



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

작물학박사 학위논문

벼 Cyanidin 3-Glucoside 함량 관여
양적형질유전자좌 분석

Identification of QTLs for High-Dense Cyanidin
3-Glucoside Content in Rice Grain

2017년 8월

서울대학교 대학원
식물생산과학부 작물생명공학전공
함 태 호

작물학박사 학위논문

벼 Cyanidin 3-Glucoside 함량 관여
양적형질유전자좌 분석

Identification of QTLs for High-Dense Cyanidin
3-Glucoside Content in Rice Grain

2017년 8월

서울대학교 대학원
식물생산과학부 작물생명공학전공
함 태 호

벼 Cyanidin 3-Glucoside 함량 관여 양적형질유전자좌 분석

지도교수 고 희 중

이 논문을 작물학 박사 학위논문으로 제출함
2017년 4월

서울대학교 대학원
식물생산과학부 작물생명과학 전공
함 태 호

함태호의 박사학위논문을 인준함
2017년 6월

위 원 장	이 변 우	(인)
부위원장	고 희 중	(인)
위 원	서 학 수	(인)
위 원	김 도 순	(인)
위 원	류 수 노	(인)

국 문 초 록

C3G(Cyanidin 3-Glucoside)는 검정색 종피 유색미의 주요색소인 안토시아닌을 구성하는 주요성분이며 항산화능력이 우수하고 흑자색미 색소의 주 구성성이다. C3G 함량에 관여하는 양적형질 유전자좌(QTL)와 후보유전자를 찾기 위해 연구를 수행하였다. C3GHi 품종에 흑진주와 수원425를 부분으로 하는 교배조합을 만들어 Backcrossed RILs 유전분석 집단을 육성하였다. 각 조합은 F₂ 세대에서 SSD(Single Seed Descent) 방법으로 F₅ 세대까지 세대진전을 진행하였고 F₅ 세대 RILs 집단에서 각 개체별로 기초농업형질인 간장, 수장, 수수를 조사하고 수확하여 C3G 함량을 측정하였다. 집단의 개체수와 C3G함량의 분포도를 비교하여 C3GHi/Suwon 425 집단을 분석집단으로 사용하였다. 유전분석 방법은 NGS(Next Generation Sequencing) 기술을 이용한 GBS(Genotyping by Sequencing)를 이용하였다. GBS 방법을 통해 총 6,703개의 SNPs가 탐색되었다. 이중 무의미하거나 중첩되는 SNP 마커들을 제외하는 선별 작업을 거쳐 475개의 의미 있는 SNP 마커를 선별하였다. 선별된 475개의 SNP 마커의 염기서열 위치를 이용하여 물리지도(Physical map)를 작성하고 SNP 마커의 유전자형을 토대로 연관지도(Linkage map)를 구성하였다. 총 17개의 연관군(Linkage group)이 구성되었는데 물리지도상에 gap이 컸던 4번 염색체는 3개 그룹으로 나누어졌고 1번, 5번, 11번 염색체는 2개의 그룹으로 나누어졌다. 연관지도에는 473개의 마커가 위치하였고, 마커간 평균 거리는 3.55cM 이었다. SNP 마커와 C3G함량간의 연관성평가(association analysis)를 수행하기 위해 분산분석(Single-effect ANOVA)을 실시하여 2번, 5번, 8번, 9번, 10번, 12번 염색체에서 C3G 함량과 연관된 6개의 QTL 위치를 찾았다. 유전자 효과가 큰 2번과 9번 염색체 QTL 내에서 후보유전자를 찾기 위해 마커를 추가하여 마커간의 거리를 좁혔다. 그 결과 2번 염색체 QTL에서 색소의 안전성에 관여하는 유

전자와 9번 염색체 QTL에서 플라보노이드 생합성 시작 아미노산인 phenylalanine 생합성 과정에 관여하는 유전자를 발견하였다.

통상 흑미로 이용되고 있는 흑진주와 초형이 좋은 흑자미 수원 425를 이용하여 C3G함량이 교배모본보다 8배가량 높은 중간모본 C3GHi를 육성하였다. 이를 이용하여 C3G 함량과 관련된 유전자를 찾기 위한 집단을 만들어 GBS방법을 통해 발굴된 SNP마커를 이용하여 간장, 수장, 수수 그리고 C3G함량 관련 QTL을 탐색하였다. 이와 동시에 C3GHi를 중간모본으로 하는 전통적인 교배육종을 실시하였다. C3G함량이 높은 C3GHi계통에 백색 종피를 가진 대립변 1호를 교배하고 계통을 전개하였다. 계통의 초형과 C3G함량을 기준으로 계통선발로 육성된 계통을 KNOU6로 하고 품종명을 슈퍼자미2호로 하였다. 천연색소(C3G) 함량을 높인 슈퍼자미2호(흑진주벼/수원 425호//대립변 1호) 품종의 중부평야지에서 평균출수기는 8월 30일로 만생종이며 모본인 흑진주보다 1달가량 늦다. 간장은 106 cm 정도로 큰편이며 임실률은 82.0%였다. 또한 현미천립중은 30.1 g 정도로 일반 흑미에 비해 무거운 편이며, 슈퍼자미2호 품종의 현미 장폭비는 2.08의 장원형으로 흑진주벼에 비해 길이와 폭이 12%, 26%로 증가된 대립종이다. 미량원소 중 K, Ca 함량은 흑진주벼에 비해 낮았고, 단백질, 회분 함량도 낮게, 열량, 지방, 탄수화물 함량은 비슷한 수준으로 나타났다. 슈퍼자미2호 품종의 C3G 색소함량은 2013년에는 1,782 mg(100g 종자), 2014년에는 1,980 mg(100g 종자)으로서 슈퍼자미 보다는 다소 낮았으나 흑진주벼 보다는 9배 이상 높았다. 슈퍼자미2호는 고기능성 품종으로 직접 이용되거나 다른 기능성 품종을 육성하는 데에 활용될 것으로 기대된다.

주요어 : 유색미, 안토시아닌, GBS(Genotyping by Sequencing), 연관지도(Linkage mapping), 양적형질 유전자좌(QTL)

학 번 : 2003-30358

목 차

I. 서론	1
II. 연구사	4
III. 재료 및 방법	17
1. 분석집단육성	17
2. 여교배집단 유전분석	19
3. 품종육성 계통재배 및 작물학적 특성조사	22
4. C3G 성분 함량 분석 및 일반성분 분석	23
IV. 결과	27
1. GBS기술을 이용한 C3G 함량관련 후보유전자·마커 동정	27
1) 농업형질 및 C3G 함량 분석	27
2) GBS를 이용한 genotyping 및 고밀도 지도 작성	29
3) 농업형질 및 C3G 함량 관련 QTL 분석	34
4) Single Effect ANOVA 분석을 통한 연관부위 탐색	42
5) C3G 함량 관련 후보유전자 haplotype 분석	58
2. C3G 고함유 기능성 품종 육성	61
1) 출수기 및 주요 작물학적 특성	61
2) 도정 및 품질특성	62
V. 결론	64
VI. 종합고찰	66
VII. 참고문헌	69
VIII. 부록	85

표 목 차

표 1. 교배모본의 주요 작물학적 특성	18
표 2. 교배조합과 후대계통수	18
표 3. NGS 라이브러리 품질 결과	29
표 4. SNP마커를 이용한 연관지도	33
표 5. C3G 함량과 간장, 수장, 수수의 양적형질유전자위치 . .	36
표 6. 간장형질 관련 QTL의 분산분석결과	43
표 7. 수장형질 관련 QTL의 분산분석결과	44
표 8. 수수형질 관련 QTL의 분산분석결과	45
표 9. C3G함량 관련 QTL의 분산분석결과	46
표 10. 슈퍼자미2호의 주요 작물학적특성	61
표 11. 슈퍼자미2호의 도정 및 품질특성	62
표 12. 슈퍼자미2호의 일반성분 함량	62

그림 목차

그림 1. C3GHi 품종과 모본 부분 품종 흑진주, 수원 425	18
그림 2. GBS 분석을 위한 라이브러리 퀄리티 테스트 결과	19
그림 3. 슈퍼자미2호 품종 육성계보도	23
그림 4. C3GHi/흑진주와 C3GHi/수원 425 교배집단의 C3G 함량 분포도 .	27
그림 5. 교배후대 집단 주요 작물학적 특성의 분포도	28
그림 6. GBS로 생성된 6,703개 SNP를 이용한 물리지도	29
그림 7. SNP 마커 필터링 과정	30
그림 8. 최종 475개 SNP 마커를 이용한 분석집단의 genotyping . . .	31
그림 9. 475개 SNP 마커를 이용한 물리지도	32
그림 10. 연관지도와 간장, 수장, 수수 그리고 C3G함량 QTL	34
그림 11. 간장관련 QTL 분석결과(염색체 6, 염색체 7)	37
그림 12. 간장관련 QTL 분석결과(염색체 6, 염색체 7, 염색체 9) . . .	38
그림 13. 수장관련 QTL 분석결과	39
그림 14. 수수관련 QTL 분석결과	40
그림 15. C3G함량관련 QTL 분석결과	41
그림 16. C3G함량 QTL갯수와 계통의 C3G함량 비교	48
그림 17. C3G함량관련 2번염색체 QTL 거리 좁히기	49
그림 18. C3G함량관련 5번염색체 QTL 거리 좁히기	50
그림 19. C3G함량관련 8번염색체 QTL 거리 좁히기	51
그림 20. C3G함량관련 9번염색체 QTL 거리 좁히기	52
그림 21. C3G함량관련 10번염색체 QTL 거리 좁히기	53
그림 22. C3G함량관련 12번염색체 QTL 거리 좁히기	54
그림 23. C3G함량관련 2번 염색체 QTL부위 유전자 탐색	55
그림 24. C3G함량관련 2번 염색체 QTL부위 CAPS 마커제작	56
그림 25. C3G함량관련 9번 염색체 QTL부위 CAPS 마커제작	57
그림 26. 상용품종 qc3g2 위치 Haplotype 분석	58
그림 27. 상용품종 qc3g9 위치 Haplotype 분석	59
그림 28. 슈퍼자미2호 C3G 함량 비교	63

I. 서론

산업발전에 따른 재배면적의 감소, 식생활 수준의 향상에 따른 음식 섭취의 변화 등으로 쌀 생산량 및 소비량은 급격히 줄어들고 있다. 특히 서구화된 음식섭취 변화는 당뇨, 고혈압, 비만 등의 만성대사성질환을 일으키는 주원인으로 지목되고 있다. 우리나라에서는 한 사람당 연평균 쌀 소비량이 지난 1979년의 135.6 kg을 고비로 점차 줄어들어 2016년에 61.9 kg 수준으로 떨어졌고, 더불어 쌀의 소비성향이 고급화, 다양화됨에 따라 소비자들의 유색미, 향미 등 기능성 쌀 품종에 대한 요구도가 높아지고 있다. 과거의 생산량만을 위한 육종, 생산자 위주의 품종 개선에서 최근의 벼 육종은 쌀 소비를 촉진하고 소비자의 요구에 부응하기 위한 품질향상은 물론 기능성 강화, 가공적성, 특수용도 품종 육성 또한 중요한 목표로 자리매김하고 있다(Ham *et al.* 2015, Ryu 2014).

유색미의 색소는 호분층과 종피, 과피를 포함한 미강에 대부분 존재한다. 유색미의 미강은 색소뿐 아니라 오리자놀, 피틴산, 페루릭산, 헤미셀룰로오스, 토코페롤, 옥타코사놀 등을 함유하고 있다(Kim *et al.* 2011).

벼 종피의 색소는 흑자색계열의 안토시아닌계 색소와 적갈색계열의 탄닌계 색소가 보고되었다. 탄닌 계열은 카테킨(catechin)과 에피카테킨(epicatechin)이 폴리머 형태로 존재하며(Finochchiaro *et al.* 2007; Mohanla *et al.* 2012), 안토시아닌 계열은 cyanidin 3-glucoside(C3G),

peonidin 3-glucoside(P3G)가 주요 구성성분이며 cyanidin 3-rhamnoglucoside, malvidin 3-galactoside 등이 존재한다(Reddy *et al.* 1994).

이러한 종피색에 관련된 유전자 위치는 흑자색계열은 *Prp b*(*Pb*, Chr. 4)와 *Prp a*(*Pp*, Chr. 1)로 알려져 있으며 적갈색계열은 7번 염색체와 1번 염색체에 각각 위치하는 *Rc*와 *Rd*로 알려져 있다(Nagao and Takahashi 1956). 이러한 종피색 관련 유전자 위치는 서로 보족적으로 작용하며(Hsieh and Chang 1964), 후대에 다양한 색깔 변이가 표현된다.

표현형질에 관여하는 유전자를 찾기 위하여 다양한 방법들이 사용되었고 근래에는 분자마커를 이용한 방법이 각광을 받고 있다. 최근에는 NGS 방법을 통한 분석 방법이 대량의 샘플에서 단시간내에 유효한 마커를 발굴하는 방법으로 많이 사용되고 있으며 특히 비용이 저렴하며 유용한 대량의 마커 탐색이 가능한 Genotyping by Sequencing(GBS) 방법은 집단의 유전 분석이나 QTL 탐색은 물론(Furuta *et al.* 2017) 다른 방법들에 비해 상대적으로 밀도 높은 마커 데이터를 생성함으로 GWAS 분석에도 적합한 방법으로 알려져 있다(He *et al.* 2014; Sonah *et al.* 2015; Begum *et al.* 2015).

본 논문에서는 벼 종피에 축적되는 색소의 함량이 연속변이를 보이며 교배 모본의 함량 보다 5~10배 높은 후대 계통이 분리되는 현상이 1~2개의 유전자에 의한 단순한 유전형상이 아니라 여러 유전자에 의해 영향을

받는 양적형질임을 가정하고 이를 구명하기 위하여 흑자색품종끼리 교배하여 분석집단을 만들었다. 안토시아닌 색소의 구성성분인 C3G 고함유 중간계통 C3GHi와 계통의 모본을 여교배한 후대 집단을 이용하여 종피의 색 유무와 달리 색소의 함량이 많고 적음에 관여하는 양적형질 유전자좌 (QTL)를 GBS 방법으로 생성한 마커를 이용하여 탐색하고 해당위치에 존재하는 유전자를 찾기 위한 연구를 수행하였다. 또한 C3G 고함유 중간모본을 이용하여 C3G 함량은 유지하며 초형과 수량성을 개선하기 위한 품종육성을 수행하였다.

II. 연구사

최근 평균수명의 증가로 선진국들은 물론 개발도상국들도 급격 속도로 노령화사회로 진입하고 있다. 이로 인해 노화와 관련된 각종 퇴행성 신경 질환과 심혈관계 질환등의 질병에 관심이 커지고 있으며 그 원인은 활성 산소에 기인한다는 보고가 있다(Jayaprakasha *et al.* 2001; Fang *et al.* 2002; Kuhn 2003). 호흡을 통해 들어온 산소는 생명유지를 위해 반드시 필요하지만 약 2~3% 정도는 체내에서 매우 불안정한 물질로 변화하여 독성을 갖는다. 특히 활성산소종(reactive oxygen species, ROS)은 세포의 주요 성분인 단백질, 핵산, 지질 등에 손상을 유발하여 세포기능에 큰 장애를 미친다(Gordon 1996; Temple 2000). 몸을 구성하는 물질들이 활성 산소에 의해 산화되는 것을 산화스트레스(oxidative stress)라 하는데 여기에서 발생하는 O_2^- (superoxide radical), OH^\cdot (hydroxyl radical), H_2O_2 (hydrogen peroxide), NO(nitric oxide)와 같이 전자가 쌍을 이루지 못하고 활동하는 free radical 및 lipid peroxide radical이 직접 관여하며, O_2^- 활성산소종이 가장 많이 생성된다. 현대인은 과도한 인스턴트식품 섭취, 심한 운동, 흡연, 심리적 스트레스의 증가와 환경호르몬이나 배기가스 등의 화학물질에 자주 노출되면서 체내 활성산소의 생성이 증가하고 있으며 이러한 산화스트레스는 암, 천식, 심혈관계 질환, 당뇨병, 염증반응 등 다양한 질병 영향을 미친다고 밝혀졌다(Nordmann 1993; Lee *et al.*

2006).

이러한 질병들이 식생활과 밀접한 관련이 있는 것으로 밝혀지면서 식품 중에 함유된 생리활성물질에 대한 연구가 다양한 식품이나 재배식물 그리고 녹차 등을 이용하여 이루어지고 있다. 생리활성 성분들 중 flavonoids와 같은 물질들은 특히 암이나 관상동맥 질환 예방에 유효하다는 것이 증명되면서(Hollman *et al.* 1996; Kenkt *et al.* 2002) 생리활성 특성을 강조한 건강 기능성식품을 이용한 질병예방 및 영양상태 개선에 사람들의 관심이 고조되고 있다. flavonoids는 2개의 기본적인 벤젠 고리 구조에 -OH기의 이중결합, 케톤그룹의 유무 등에 따라서 크게 flavonol과 flavone(anthoxanthin), flavanoneol, flavan, anthocyanidin, isoflavone 등으로 구분되며(Winkel-Shirley 2001), free radical을 효과적으로 제거함으로써 항염증(Min *et al.* 2010), 저혈당(Tsuda *et al.* 2003), 항암(Hyun and Chung 2004) 등의 작용을 보인다고 밝혀졌다(Di Carlo *et al.* 1999).

쌀의 현미층은 식이섬유를 다량 함유하고 있으며 각종 단백질, 불포화지방산 및 비타민, 무기질 등이 풍부하여 영양학적으로 우수하고 γ -oryzanol, phenolic acid, ferulic acid, phytic acid 과 oryzafuran 등 같은 다양한 유용성분을 포함하고 있다(Zhang *et al.* 2005). 페놀성 화합물은 식물계에 널리 분포되어 존재하는 2차 대사산물 중 하나로 다양한 구조와 분자량을 가지며 phenolic hydroxyl기가 단백질과 같은 거대

분자와 결합하여 항산화, 항암 및 항균 등의 생리기능을 가지는 것으로 알려져 있다(Choi *et al.* 2003). 유색 현미는 적갈색, 흑자색 등의 다양한 천연색소를 함유하는데 일반적으로 흑자색 유색미에는 안토시아닌계 색소가 다량 함유되어 있고 적갈색 유색미에는 탄닌계 색소가 함유되어 있는 것으로 알려져 있다. 안토시아닌계 색소는 높은 항산화 기능(Osawa 1992; Han *et al.* 2004; Kim *et al.* 2010) 뿐만 아니라 돌연변이 억제효과 등의 높은 기능성으로 인체 면역력을 향상시키며 항암(Dai J. *et al.* 2009), 항아토피(Han *et al.* 2007; Han *et al.* 2009), 항당뇨(Kim *et al.* 2010) 및 심혈관계 질병 등의 예방과 치료효과(Verlangieri *et al.* 1985) 등이 연구·보고 되면서 기능성 소재의 원료로서 각광을 받고 있다. 탄닌계 색소 역시 유해성 중금속을 제거하고 변이원 생성을 억제시키는 것으로 알려져 있다(Oh *et al.* 1996; Ryu 2014). 흑자색미 색소인 안토시아닌 색소는 cyanidin 3-glucoside(C3G)가 80% 이상을 차지하며 그 밖에 peonidin-3glucoside(P3G), cyanidin 3-rhamnoside, malvidin 3-galactoside, pelargonidin 및 selphinidin 등으로 구성되어 있다(Terahara *et al.*, 1994; Reddy *et al.* 1995; Ryu *et al.* 1998; Park *et al.* 2008).

식물의 이차대사물질은 다양한 환경적인 재해, 곤충, 박테리아, 곰팡이 같은 병원성에 내성을 가지게 된다(Harborn and Williams 2000; Winkel-Shirley 2001). 다양한 이차대사물질 중 플라보노이드는 가장 많

이 연구가 이루어진 물질의 하나로서 현재 까지 약 10,000여종 이상의 식물에 존재하는 것으로 알려져 있다(Tahara 2007). 플라보노이드는 나비와 벌과 같은 곤충을 유인하여 꽃의 수정을 돕기 위하여 꽃의 색깔을 내거나 과일의 색소를 내는데 관여하기도 하고, 질소고정균과 같은 토양박테리아 및 식물호르몬의 신호전달물질로도 사용된다(Harborn and Williams 2000; Williams 2000; Winkel-Shirley 2001). 또한 대기권에 들어오는 자외선 차단 및 식물의 병원성 균으로부터 식물을 보호해주는 파이토알렉신(phytoalexin)과 타감작용(allopathic compound)을 한다(Williams 2000; Winkel-Shirley 2001). 최근 들어 발표되는 다양한 연구의 결과에서 알 수 있듯이 인간의 건강에 이로운 항암, 항알러지, 항바이러스, 심장병예방, 항산화 작용, 골다공증예방과 같은 다양한 생리적 효과를 가지고 있는 것으로 알려져 있어 새로운 신약을 개발하기 위한 잠재적인 요소를 가지고 있기 때문에 이들 물질들에 대한 생리적 역할에 대한 많은 연구가 이루어지고 있다(Harborne and Williams 2000; Ross and Kasum 2002). 모든 플라보노이드는 phenylbenzopyrone의 기본 골격을 가지고 있다. 플라보노이드 합성의 시작은 4CL(4-coumarate:CoA ligase)에 의해서 P-coumaroyl-coA가 만들어지고 CHS(chalcone synthase)는 이렇게 만들어진 P-coumaroyl-coA에 세분자의 malonyl-CoA가 순차적으로 결합하여 chalcone naringenin을 생성하게 된다. chalcone naringenin은 CHI(chalcone isomerase)에 의해서 isomerization이 되어 2S-나린제

닌을 생성하게 된다(Schijlen *et al.* 2004). 2S-나린제닌은 플라보노이드
 생합성에 필요한 다양한 효소들의 전구물질로 사용되어진다. 2S-나린제닌
 은 *FNS(flavone synthase)*에 의해서 oxidation이 일어나 아피제닌을 생
 성하게 된다. *FNS* 유전자는 크게 *FNS I* 과 *FNS II*의 두 그룹으로 나눌
 수 있다(Heller and Forkmann, 2003). *FNS I*은 2-oxoglutarate-/
 Fe^{2+} dependent dioxygenase 유전자로서 주로 수용성단백질로 존재한다.
 이에 반해 *FNS II* 효소는 cytochrome P450에 속하는 monooxygenase
 로서 이것은 membrane binding 단백질로서 NADPH가 필요로 한데 이것
 은 P450 reductase에 의해서 전자를 공급받아 사용한다(Heller and
 Forkmann, 2003). F3H(flavanone 3 β -hydroxylase)는 2S-나린제닌을
 기질로 하여 (2R, 3R)-trans-dihydroflavonols을 형성하게 되고, 이것은
 다시 FLS(flavonol synthase)에 의해서 oxidation이 일어나 플라보놀이
 생성한다. 이와 같이 만들어진 flavonol은 당전이효소
 (glycosyltransferase), 메틸전이효소(methyltransferase), 프레닐전이효소
 (prenyltransferase), F3'H(flavonoid 3'-hydroxylase), F3'5'H(flavonoid
 3',5'-hydroxylase) 등에 의해서 다양한 물질들을 생성하게 된다
 (Winkel-Shirley 2001). FHT와 FLS 두 효소는 *FNS I* 효소와 마찬가지로
 반응중에 2-oxoglutarate를 필요로 하는 dioxygenase이다(Heller
 and Forkmann 1994). 또한 F3H에 의해서 만들어진 dihydroflavonol을
 기질로 하여 DFR(dihydroflavonol 4-reductase)에 의해서

leucoanthocyanidin을 생성하게 된다. LDOX(leucoanthocyanidin dioxygenase)는 DFR에 의해서 생성된 leucoanthocyanidin을 축매하여 anthocyanidin을 생성하게 된다. Anthocyanidin은 3GT(anthocyanidin 3-O-glycosyltransferase)에 의해서 anthocyanin을 합성한다. 또한 anthocyanidin은 ANR(anthocyanidin reductase)에 의해서 catechin(flavan 3-ols)을 형성하게 되는데 이것은 프로안토시아닌 (proanthocyanidin) 합성의 전구 물질로 사용된다(Owens *et al.* 2008). 플라보노이드 생합성에 있어 위와 같은 여러 효소들이 중요한 역할을 담당하고 있지만, 이와 더불어 전사조절인자 또한 이들의 생합성에 있어 아주 중요한 역할을 하고 있다. 옥수수에서 발견된 *C1*(R2R3 MYB 전사조절인자), *R*(basic helix-loop-helix-type)전사조절인자는 독립적으로 또는 상호 의존적으로 플라보노이드 생합성을 조절하는 것으로 알려져 있다(Quattrocchio *et al.* 1993, 1998). *R2R3*와 *bHLH* 유전자들은 많은 쌍자엽식물과 단자엽식물에서 분리되었으며, 이러한 유전자들간의 유사성은 상당히 높은 편이다. 이러한 결과는 이들 유전자들이 다른 식물체에 발현시켰을 경우에도 동일한 기능을 할 수 있다는 것을 암시한다(Schijlen *et al.* 2004). 실질적인 예로 애기장대와 담배에 옥수수에서 분리한 *C1*과 *R*을 과발현 시켰을 때 뿌리와 엽병, 줄기조직이 애기장대와 담배의 야생형과 비교했을 때 안토시아닌이 현저하게 증가하는 것을 보였다(Lloyd *et al.* 1992). 또한 옥수수의 *C1*과 *LC*(옥수수 R유전자 family)를 토마토에

과발현 시켰을 때 토마토의 과육조직과 잎에 안토시아닌과 플라보노이드의 함량이 현저하게 증가되었다(Bovy *et al.* 2002). 이렇게 플라보노이드와 안토시아닌의 함량이 증가된 것은 플라보노이드 생합성에 관여하는 *CHS*, *CHI*, *FLS*와 같은 유전자들의 전사량이 증가하였기 때문이다(Verhoeven *et al.* 2002). 옥수수의 *CI*과 *R-S* 유전자를 벼에서 endosperm specific 프로모터를 가지고 과발현 시켰을 때 벼의 종실에 taxifolin, dihydroisorhamnetin, isorhamnetin과 같은 다양한 플라보노이드가 집적되었다(Shin *et al.* 2006). 애기장대의 *PAP1*(PRODUCTION OF ANTHOCYANIN PIGMENT1, *AtMYB75*)과 *PAP2* 유전자는 activation tagging에서 의해서 분리되었는데, 이들 유전자들은 안토시아닌과 quercetin glycoside의 생합성을 증가시켰다(Borevitz *et al.*, 2000). 그 결과 애기장대의 붉은 보랏빛의 색깔의 잎을 보였고, 뿌리와 꽃은 핑크색으로 변하였다. 이것은 *PAP1*, *PAP2* 전사조절인자에 의해서 플라보노이드 생합성에 관여하는 유전자, *PAL*, *CHS*, *CHI*, *DFR*, *ANS*와 같은 유전자들의 전사가 증가되었기 때문이었다(Boreriz *et al.* 2000; Matsui *et al.* 2004; Tohge *et al.* 2005).

벼 종피 흑자색에 관련된 유전자좌는 4번 염색체 위치한 *PURPLE PERICARP B*(*Prp-b*, *Pb*)와 1번 염색체에 위치한 *PURPLE PERICARP A*(*Prp-a*, *Pp*)로 알려져 있고, 적갈색은 1번 염색체에 위치한 *Rd*와 7번

염색체에 위치한 *Rc* loci로 알려져 있다(Nagao and Takahashi 1947). 흑자색과 적갈색에 관련된 loci *Pb*, *Pp* 와 *Rc*, *Rd*는 모두 보족적으로 작용하여(Rahman *et al.* 2013; Wang and Shu 2007; Nagao and Takahashi 1947) *Pp*와 *Rd*는 색소 축적에 관여 하고, *Pp*와 *Rc*는 색소의 축적정도를 관여하는 것으로 보고되었다. 예로, *Pb_PpPp* 유전자형은 진한 흑자색 종피로 표현되며 *Pb_Pppp* 유전자형을 갖는 계통 혹은 품종은 중간 흑자색 종피로 표현된다. 그러나 *Pb_pppp* 유전자형은 갈색으로 표현되고 무색 종피인 *pbpbPp*-와 *pbpbpppp* 유전자형과 마찬가지로 안토시아닌이 측정되지 않았다(Rahman *et al.* 2013). Furukawa *et al.*(2006)은 *Rd* 위치에 *DFR*(*Dihydroflavonol reductase*) 유전자가 존재한다고 보고 하였으며 Sweeney *et al.*(2006)은 *Rc* 위치에 *bHLH*(*basic Helix-loop-helix*) 발현 유전자를 보고하였다. 식물의 여러 조직에서 색소 침착은 *Chromogen* 유전자(*C*), *Activator* 유전자(*A*) 그리고 tissue-specific 조절 유전자(*P*)가 필요하며, *C*는 6번 염색체에 위치하며 6개의 대립유전자가, *A*는 4번 염색체에 위치하며 4개의 대립유전자가 보고되었고, 조절유전자 *P*는 1번 염색체에 위치한다고 보고되었다(Nagao and Takahashi 1963; Kinoshita 1995). Koes *et al.*(2005) 는 안토시아닌과 프로안토시아니딘 색소의 침착 조절에 관련하여 MYB, HLH 그리고 WD40 transcription factor가 *DFR* 유전자의 발현에 관여하며 이를 통해 각 조직의 색소 침착이 조절된다는 모델을 제시하였다.

MBW(MYB-bHLH-WDR) 단백질 복합체는 flavonoid 생합성과정을 환경적요인에 의한 미세한 전사조절부터 생합성 시작에 관여하는 positive 조절에 까지 관여하는 생합성 조절 유형으로 보고되었다(Xu *et al.* 2015)

유전자지도 작성 기술은 QTL 등을 이용하여 육종 효율을 증가시키기 위한 기술로 병해충 저항성 있는 유전자의 집적, QTL을 함유하는 작물을 선발하는 방법, 양적형질을 결정하기 위한 genetic marker를 사용하는 방법, 복합 형질에 대한 QTL을 확인하는 방법 등의 기술이 개발되어 활용되고 있다. 교배·선발 방식 육종법을 이용하여 우수 품종을 육성할 때 분자마커와 같은 첨단 육종기술을 이용하면 육성기간 단축뿐만 아니라 조기선발을 통해 작물의 유지 및 관리에 소요되는 많은 경비를 줄일 수 있다. 이와 같은 실질적인 효과가 입증되면서 국외뿐만 아니라 국내 육종 현장에도 분자마커 활용 요구도가 급증하고 있는 추세다.

유전형상을 밝히기 위한 유전 마커는 이미 멘델의 실험에서부터 시작되었고 그 후 초파리를 이용한 표현형 마커는 genetic linkage 뿐 아니라 연관지도 작성의 효시가 되었다. 여러 표현형 마커들의 한계성으로 DNA에 직접적으로 작용하는 마커를 개발하게 되었고, 분자마커는 지난 수십 년동안 끊임없이 발전해 왔다(Agarwal *et al.* 2008). 대표적인 종류는 처음으로 분자마커로 개발된 RFLP(Restriction Fragment Length Polymorphism: Botstein *et al.* 1980)가 있다. 제한효소를 이용하여

labeling된 DNA probe를 이용하여 다형성을 찾는 방법의 마커이다. 다양한 제한효소를 이용함으로써 전체 게놈에 걸친 마커를 찾을 수 있고, 제한효소 특성상 특정 염기서열을 자르기에 염기서열 사이에 삽입되거나 삭제된 염기서열 혹은 SNP에 특화되었다. 그러나 시간과 돈이 많이 들고 방사선, 독성시약 등을 사용해야 하며 DNA 순도가 높아야 하는 단점들로 인해 오랫동안 사용되지는 않았다. 이후 Mullis and Faloona(1987)의 PCR 기술개발로 PCR을 기반으로 한 다양한 분자마커가 개발되었다. 10-mer bases를 사용한 random primer를 이용한 RAPD(Randomly Amplified Polymorphic DNA: Williams *et al.* 1991)마커는 빠르고 효율적이라는 장점을 가지고 많은 작물에서 고밀도 지도 작성에 사용되어 알팔파(Kiss *et al.* 1993), 참두(Torress *et al.* 1993), 사과(Hemmat *et al.* 1994)의 지도 작성에 사용되었다. 또한 토마토, 상추, 강낭콩 등에서는 RAPD 분석을 통해 병저항성 유전자를 동정하기도 하였다(Martin *et al.* 1991; Paran *et al.* 1991; Adam-Blondon *et al.* 1994). 하지만 PCR반응 조건에 따라 loci가 달라지거나 heterozygous 판별이 어려운 단점이 있다(Bardakci 2001). 이러한 RAPD마커의 재현성을 개선하기 위해 AFLP(Amplified Fragment Length Polymorphism: Vos *et al.* 1995) 마커가 개발되었다. RFLP의 장점과 RAPD의 장점을 혼합하였다. 제한효소로 잘린 DNA 단편에 인식이 가능한 염기서열을 adaptor로 붙이고 여기에 특정 primer를 사용하여 DNA 단편을 증폭시키는 방법이다. 또한

형광 프라이머등을 사용하여 자동화도 가능하다(Huang and Sun 1999).

SSR(Simple Sequence Repeat: Litt and Luty 1986; Wenz *et al.* 1998)은 흔히 microsatellite 마커로도 불리는 진핵세포에 넓게 산재해 있는 단조롭고 짧게 반복되는 염기서열을 이용한 분자마커이다(Tautz and Renz 1984). 단순반복서열은 DNA 복제중 strand slippage에 의해 생김(Schlotterer and Tautz 1992), point mutation 보다 일어나기가 쉽기 때문에 단순반복서열은 종간 혹은 종내에서도 다형성을 가진다. 공우성으로 유전되고 대립유전자의 다양성도 풍부하며 사이즈 평가가 쉽다는 장점을 갖고 있고 재현성도 뛰어나서(Saghai Marroof *et al.* 1994) 상추나 보리 벼 등의 다양한 작물에서 분자마커로 쓰이고 있다(van de Wiel *et al.* 1999; Saghai Marroof *et al.* 1994; Wu and Tanksley 1993).

RAPD나 AFLP 마커 분석으로 확인된 마커를 map-based cloning에 유용하게 쓰기위하여 SCAR(Sequence characterized amplified region: Paran and Michelmore 1993; McDermott *et al.* 1994)마커가 개발되었다. PCR 생산물 양끝의 염기서열을 분석하고 이를 이용하여 특이적인 프라이머를 제작하는 방법으로 유전체 풀의 스크리닝뿐 아니라 물리지도 작성이나(Chelkowski and Stephen 2001), 특정 유전자 위치를 확인하거나(Paran and Michelmore 1993), comparative mapping(Guo *et al.* 2003)에 사용되었다. CAPS(Cleaved Amplified Polymorphic Sequences: Konieczny and Ausubel 1993)마커는 RFLP 마커 위치의

염기서열정보를 PCR-based 마커로 전환시킨 것이다. SNP나 In/del 등에 의해 나타나는 길이의 다형성을 이용한다(Chelkowski and Stephen 2001; Konieczny and Ausubel 1993). SNP(Single Nucleotide Polymorphism)는 genome 전체에 걸쳐 가장 풍부하게 존재하여 광범위한 genome 분석에 적합하며(Rafalski 2002; Zhu *et al.* 2003), 최근에는 차세대 유전자염기서열분석(NGS : Next Generation Sequencing)을 이용한 대량 유전체 정보를 활용하는 방법들이 개발되어 가격적인 면이나 시간적인 면에서도 장점을 갖추게 되었다(Davey *et al.* 2011).

SNP 마커는 다른 마커들에 비해 유전체 내 발생빈도가 가장 빈번하여 옥수수에서는 60-120 염기당 1개(Ching *et al.* 2002), 인간은 1,000 염기서당 1개꼴(Sachidanandam *et al.* 2001)로 발생한다. 게다가 High-throughput 기술에 보다 적합한 특성, 즉 탐색에서 분석에 이르는 전 과정이 자동화 될 수 있다는 점이 다른 마커들보다 품종판별이나 고밀도 유전자지도 작성 등에 활용성을 높이고 있다(Agarwal *et al.* 2008). NGS를 이용하여 SNP를 탐지하는 방법은 크게 유전체 전체를 읽는 Resequencing 방식(WGS : Whole Genome Sequencing)과 대상영역을 좁혀서 씨퀀싱하는 방법(GBS, RAD-seq 등)이 있다. WGS 방식은 우수 자원이나 원종계통에 보편적으로 사용되고 있다. 유전체 모든 영역에 존재하는 SNP를 모두 확보할 수 있는 장점이 있는 반면, 유전체 전체를 씨퀀싱 해야 한다는 비용적 부담이 있다. 또한 과도하게 많은 SNP 중 실제적

으로 사용 가능한 SNP를 선발하는 일도 상당한 어려움을 일으키고 있다. 이로 인해 SNP 수를 적절한 수준으로 줄이면서 더 많은 수의 샘플이나 계통을 분석하고자 하는 요구가 발생하였다. 이에 비해 GBS(Genotyping-by-Sequencing) 기술은 Re-sequencing과 같은 NGS 기반 방식이지만 제한효소로 genomic DNA를 절단하여 절단된 주변 서열만 부분적으로 해독할 수 있는 특징을 가진다(Elshire *et al.* 2011). 유전체 전체를 시퀀싱 하지 않고 부분적으로 two-enzyme 방식(Poland *et al.* 2012) 등을 사용하여 복잡성은 줄이고 균일한 library를 생산하여 유전체 전체 정보를 확보할 수 있기에 집단유전학, QTL 탐색이나 고밀도지도 작성, genomic selection 등의 분석에 가성비가 높은 효율적인 도구로 여겨진다(Spindel *et al.* 2013; Rabbi *et al.* 2014; Huang *et al.* 2014; Burrell *et al.* 2015; Begum *et al.* 2015; Elmer *et al.* 2015; Lin *et al.* 2015).

Ⅲ. 재료 및 방법

1. 분석 집단 육성

C3G 고품량 계통인 C3GHi에 흑진주와 수원 425를 여교배하여 교배집단을 만들었다. F₂ 세대부터 SSD(Single Seed Descent) 방법으로 집단을 세대진전하여 F₅ 세대에서 개체별로 간장, 수장, 수수 및 C3G 함량 분석을 하였다.

재배방법은 중부지방에 보통기 보비재배를 하였고, 4월 28을 파종하여 5월 29일 이앙하였다. 재식거리는 30 × 15 cm로 1주 1본으로 하였다. 시비량 및 질소분시방법은 농촌진흥청 표준재배법에 준하였다.

집단 육성을 위해 쓰인 흑진주는 C3G 함량이 250mg/100 seed 이고 출수기가 7월 22일 정도인 극조생 단간 유색미 품종이며 수원 425는 C3G 함량이 150mg/100 seed 이고 출수기가 8월 16일인 중생 중장간 유색미 품종이다. C3GHi는 흑진주와 수원425의 교배육종으로 육성된 중간모본으로 키는 흑진주보다 10cm 가량 큰 단간이며 출수기는 흑진주와 수원 425의 중간이다(그림 1, 표 1). C3GHi의 C3G 함량은 1,507mg/100 seed 로 모본인 흑진주 수원 425 보다 8-10배 많다.



Heugjinju

C3GHi

Suwon425

Fig. 1. C3GHi developed from crossing between Heugjinju and S. 425

Table 1. Agronomic traits of parent

Variety	CL (cm)	PL (cm)	PN	HD	TGW (g)	C3G content (mg/100g seed)
Heugjinju	52.7±2.8	16.7±4.9	10.4±2.3	7/22	16	250
Suwon 425	77.4±4.6	21.8±1.3	11±2.4	8/16	24	150
C3GHi	55.4±3.8	21.7±1.4	11.4±3.5	8/5	17	1,507

※ CL: Culm Length, PL : Panicle Length, PN : Panicle number, HD : Heading Date, TGW : 1,000-Grain Weight

Table 2. Cross combination and No. of progeny

Cross combination	Generation	No. of progeny
Heugjinju/Suwon 425//Heugjinju	F ₅	182
Heugjinju/Suwon 425//Suwon 425	F ₅	224

분석집단 F₅ 세대에서 C3GHi(흑진주/수원 425)//흑진주 집단은 182개 계통, C3GHi(흑진주/수원 425)//수원 425 집단은 224개 계통을 얻었다.

각 집단의 농업형질은 간장과 수장, 수수를 조사하였고, 종자를 수확하여 C3G 함량을 조사하였다. C3G 함량의 변이 폭이 넓고 정규분포에 가까운 C3GHi/수원 425 집단을 QTL 분석을 위한 집단으로 선택하였다.

2. 여교배 집단 유전분석

1) DNA 추출

분석집단 각 line의 종자를 파종하여 30일 키운 후 어린잎을 개체별로 따서 2.0ml tube에 담고 액체질소에 담아 실험실로 옮긴 후 GeneJET Plant Genomic DNA Purification Mini Kit(Thermo Scientific, EU)를 이용하여 추출하였다.

2) 염기서열 해독을 위한 DNA library 제작

추출완료된 gDNA 1~5μg을 QIAquick PCR Purification kit로 정제하였다. TruSeq library sample prep kit(Illumina)을 이용하여 end repair, 3' overhang, adaptor ligation 과정을 거친 후, QIAquick PCR purification kit으로 정제하였다. Adapter가 붙은 샘플 전부를 2% Agarose gel (BioRad, USA)에 전기영동 후, 목표 크기의 밴드를 잘라내어 gel extraction kit (QIAGEN, USA)로 정제하였다. Phusion DNA polymerase를 이용하여 PCR 증폭하고 정제한 후 제한효소 *Pst*I-*Mse*I의 2-enzyme 조합으로 Truong et al.(2012) SBG 시스템을 활용하여

library 제작을 완료하였다. 제작이 완료된 library는 Bioanalyzer high sensitivity kit을 이용하여 library의 물농도를 측정하고, qPCR 방법을 이용하여 library의 정확한 물농도를 측정한 후 cluster 합성에 사용될 농도를 결정하였다.

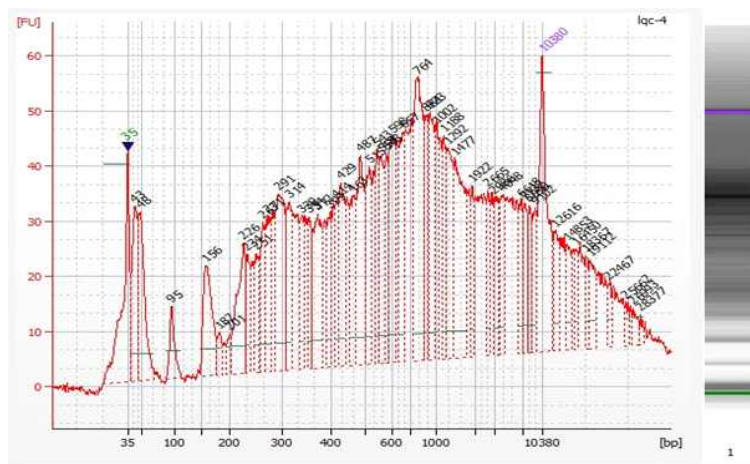


Fig. 2. Testing of Library QC for GBS analysis

3) 염기서열 해독

library 품질검사(QC)를 통과한 library는 cluster 합성을 위하여 cBot기기를 이용하여 cluster를 합성하였다. Cluster 합성이 종료되면, 염기해독 시약인 SBS kit를 원하는 염기해독 싸이클에 맞는 양만큼 준비 하고, 각각 HiSeq 2000(Illumina, USA)기기내의 정해진 위치에 로딩하였다. Cluster가 합성된 flow cell을 기기 내에 위치시키고 Reagent delivery check를 수행하여 fluidic system에 문제가 없음을 확인하였다. First

base report 프로토콜을 수행하여 cluster 생성 수를 확인하고 해독 시그널 등이 정상인지를 확인한 후 품질검사가 정상일 경우 염기해독을 진행하였다.

4) SNP 마커 필터링과 선발

Hiseq 2000으로 생산된 short read sequence를 대상으로 Tassel-GBS v.5 pipeline으로 SNP를 calling 하였다. 총 6,703개의 SNP가 calling(MAF>0.03) 되었으며 전체 12개 genome에 걸쳐 고루 분포되었다. 총 96 line 중 genotyping 데이터가 불분명한 N이 22개 이상인 SNP를 제외하여 3,434개 SNP를 필터링하여 남은 3,269 SNP를 이용하여 각 genotype 별로 2(paternal), 1(hetero), 0(maternal) 과 -1(missing)로 변환하였다. 분리가 일어나지 않은 318개 SNP 마커를 제거한 총 2,951개 SNP 마커를 QTL IciMapping 4.1 프로그램을 이용하여 최종적으로 유의한 475개 SNP 마커를 발굴하였다.

5) 고밀도 유전자 지도 작성

발굴된 475개의 SNP 마커 염기서열정보를 이용하여 물리지도(physical map)를 작성하였으며 SNP 마커 유전자형을 토대로 Joinmap 프로그램을 이용하여 연관지도(linkage map)를 구성하였다.

6) QTL 위치 유전통계 분석

SAS 프로그램(SAS Institute, 2000) PROC MEANS 이용하여 기술통계량을 산출하였고, PROC GLM을 이용하여 SNP 마커들에 의해 표지되는 유전자 위치가 간장, 수장, 수수의 농업형질과 C3G 함량 발현에 관여되는 정도를 판정하여 연관성분석(association analysis)을 수행하였다. QTL의 위치의 F 분석 결과 P 값이 0.05보다 작으면 유의하다고 판정하였고, 각 형질에 해당하는 유전자좌의 기여도와(R^2 value; %) 유전적 작용가(additive effect) 등을 추정했다.

3. 품종육성 계통재배 및 작물학적 특성 조사

흑진주와 수원425 조합 후대 중 C3G함량 뿐만 아니라 수량성을 개선하기 위하여 2003년 ‘흑진주벼’와 ‘수원425호’ 교배 후대계통 중 C3G 함량이 높은 계통(C3GHi)과 ‘대립벼 1호’를 인공교배한 F_2 세대에서 초형이 우수하고, C3G 함량이 높은 개체를 선발하였다. 2003년 동계 온실에서 세대진전 후 2007년까지 매년 계통재배하면서 포장선발을 실시하여 유전적으로 완전히 고정된 계통을 KNOU6라 명하고 추후 슈퍼자미2호로 품종명을 정하였다(그림 3). 2009년부터 2011년까지 재배시험을 통해 기본적인 농업형질인 출수기, 간장, 수장, 수수, 등숙율, 현미천립중, 제현율 등을 조사하였다. 2012년과 2013년 포장에 1주 3본, 3반복으로 이양하여 생산력 검정을 실시하였다.

Year	03	03/04	04	05	05/06	06	07	07/08	08	09	12
Generation	Cross	F ₁	F ₂	F ₃	F ₄	F ₅	F ₆	F ₇	F ₈	F ₉	F ₁₁
C3GHi × Daeribbyeo 1	SR287 21	$\left[\begin{array}{c} 1 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \end{array} \right]$ 20	$\left[\begin{array}{c} 1 \\ \cdot \\ (7) \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \end{array} \right]$ 500	$\left[\begin{array}{c} 1 \\ \cdot \\ (5) \\ \cdot \\ \cdot \\ \cdot \\ \cdot \end{array} \right]$ 50	$\left[\begin{array}{c} 1 \\ (2) \\ \cdot \\ \cdot \\ \cdot \end{array} \right]$ 50	$\left[\begin{array}{c} (1) \\ \cdot \\ \cdot \\ \cdot \\ 15 \end{array} \right]$	$\left[\begin{array}{c} 1 \\ (2) \\ \cdot \\ \cdot \\ 6 \end{array} \right]$	$\left[\begin{array}{c} (1) \\ \cdot \\ \cdot \\ \cdot \\ 6 \end{array} \right]$	$\left[\begin{array}{c} \cdot \\ \cdot \\ \cdot \\ \cdot \end{array} \right]$	$\left[\begin{array}{c} \cdot \\ \cdot \\ \cdot \\ \cdot \end{array} \right]$	RYT ^z KNO U 6

Fig. 3. Pedigree diagram of 'Superjami 2'

Z RYT : Replicated yield trial

4. C3G 성분 함량 분석 및 일반성분 분석

1) C3G 함량

종피에 함유된 안토시아닌 색소 중 C3G의 함량분석은 Ryu et al.(1998)의 방법을 따랐다. 유색미 종자를 제현 후 곱게 마쇄한 분말 2 g에 0.1% Trifluoroacetic acid (TFA)-95% EtOH 용매 20 mL를 가한 후 진탕 추출하였다. 추출액은 50 mL로 적정하여 0.45 µm PVDF syringe filter로 여과하고 HPLC (waters 501 pump, millipore gradient controller, waters 480 UV-Vis detector)를 이용하여 정성, 정량하였다. HPLC 분석은 ODS-5 (4.6 mm × 250 mm, Nomura Chemical Co., Japan) column

을 사용하였고, 530 nm 파장으로 검량하였다. 이동상으로 0.1% TFA in water, liner gradient, 0.1% in Acetonitrile을 사용하였고, 유속은 1.0 ml/min이었다.

2) 질소함량

시료에 따라 0.3g~1g을 유산지에 취하여 kjeldahal flask에 넣었다. Kjeldahal flask에 분해 촉진제 $K_2SO_4:Cu_2SO_4.H_2O(9:1)$ 와 진한 황산 (98%) 20 ml를 넣고 최대 420°C에서 1.5~2 시간 분해하였다. 실온에서 냉각 후 flask에 H_2O 40 ml과 32% NaOH 60 ml을 가하여 증류하였다. 3% Boric acid 90 ml에 NH_3 gas를 포집하여 0.1N-HCl로 적정하였다. 계산방법은 다음과 같다.

$$\text{단백질(\%)} = (A - B) * f * \text{질소계수} * 0.0014 / S * 100$$

A:0.1N-HCl 소비량(ml), B:blank test의 0.1N-HCl 소비량(ml),
f:0.1N-HCl의 역가 = 1, S:Sample 채취량(g)

3) 지 방(fat)

분말 2g 을 원통여과지에 넣고 검체 위에 탈지면을 충전하여 용기에 담은 채로 100~105°C의 건조기에서 2~2시간 건조한 후, 데시케이터에서 식히고 속슬렛추출장치의 추출관에 넣는다. 받는 그릇에 무수에테르 1/2용량을 넣어 8시간 추출한다. 추출 후 waterbath에서 에테르를 완전히 증발

시킨다. 받는 그릇을 깨끗이 닦은 후 98~100℃ 건조기에서 1시간 항량이 될 때까지 건조한 다음 데시케이터에서 식히고 칭량한다. 조지방 함량은 다음 식에 따라 산출한다(Soxxhlet extraction method 식품공전 에테르추출법).

$$\text{조지방(\%)} = (W_1 - W_0 / S) * 100$$

W_0 :받는 그릇의 무게(g), W_1 :조지방 추출하여 건조시킨 받는 그릇의 무게(g), S : 검체의 채취량(g)

4) 탄수화물(carbohydrate)

검체 100g 중에서 수분, 조단백질, 조지방 및 회분의 값을 감하여 얻은 양으로서 표시(식품공전 계산법).

5) 열 량(Calorie)

식품공전(2009) 계산법에 따라 검체 100g 중의 조단백질, 조지방 및 탄수화물 또는 당질 함량에 단백질 4, 지방 9, 탄수화물 4의 계수를 곱하고 각각의 에너지를 kcal단위로 산출하여 그 총계로 나타냈다(에트워터 계수사용법).

6) 무기성분

무기성분 중 Ca, K의 함량을 분석하기 위해 시료 1g을 각각 ICP 튜브에

담은 후 질산 5 ml을 넣고 150°C에서 습식분해법을 이용하여 전처리 하였다. ICP 분석기기를 사용하여 Ca, K 표준시약 검량곡선을 이용하여 시료의 Ca, K 함량을 분석하였다.

IV. 결과

1. GBS기술을 이용한 C3G 함량관련 유전자·마커 동정

1) 농업형질 및 C3G 함량 분석

C3GHi의 모본인 흑진주와 수원425를 각 각 여교배하여 만들어진 backcrossed population의 C3G함량과 농업형질 간장, 수장 그리고 수장은 표 1과 같다. 2개 여교배집단에서 C3G함량은 모두 연속변이를 보였으나 정규분포에서 왼쪽으로 치우친 경향을 보였다(그림 4.)

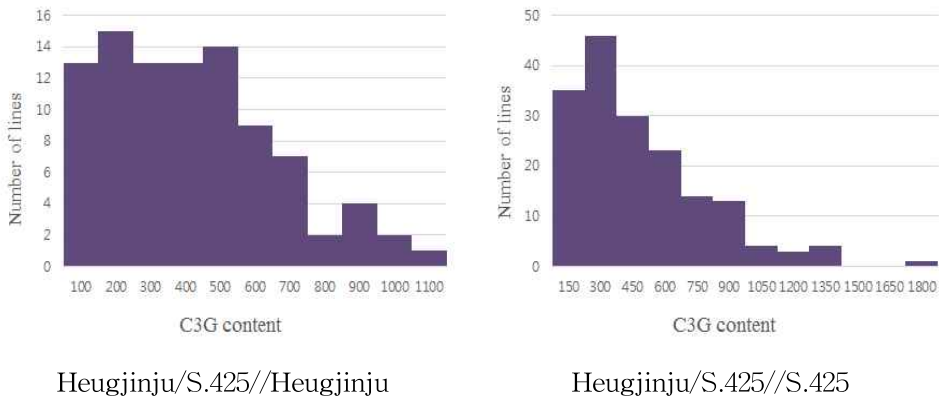


Fig. 4. Distribution of C3G content in C3GHi/Heugjinju and C3GHi/S. 425

이는 분석집단의 여교배 특성상 한쪽 부모의 영향이나 유전자형의 치우침이 생길 수 있고, C3GHi의 모본과 부분인 흑진주와 수원425가 각 각 인디카형, 자포니카 형으로 대립유전자의 구성이 서로 다르며 그 후대에

이러한 대립유전자들의 재구성으로 인해 정규분포 모습을 보이지 않는 것으로 생각된다. 특히 C3G 함량의 분포는 부모의 C3G 함량보다 10배 정도 차이가 나므로 초월분리 현상이 일어난 것으로 생각된다(Ham et al. 2015). C3G 함량에 관한 QTL 분석을 위해 C3G 함량의 변이 폭이 크고 정규분포에 가까운 C3GHi/수원 425 집단을 QTL 분석을 위한 집단으로 선택하였다. 이 집단의 간장, 수장 그리고 수수 등 농업형질을 조사한 결과는 Fig. 4와 같다. 간장, 수장 그리고 수수 형질의 분포도는 전체적으로 정규분포와 비슷한 연속변이를 보인다.

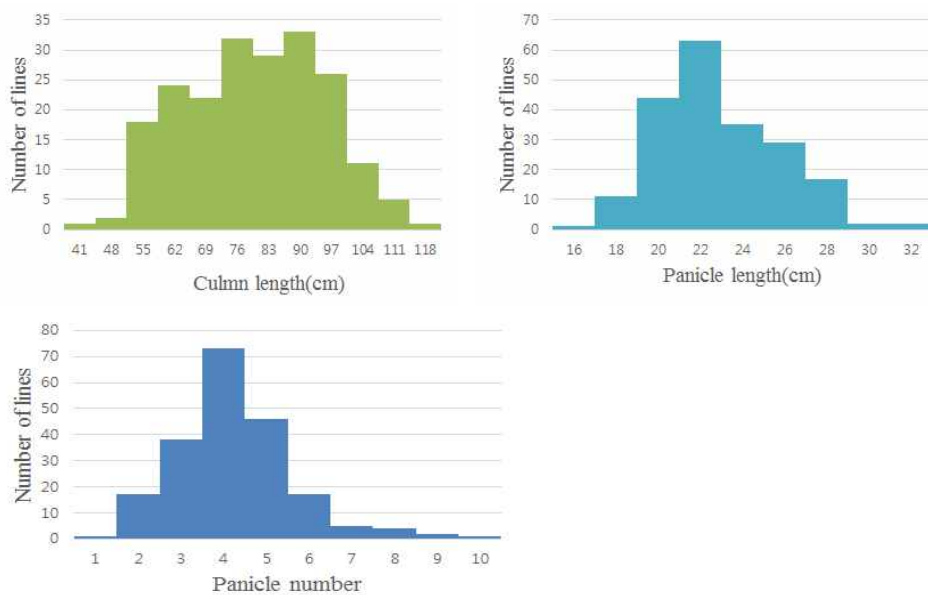


Fig. 5. Distribution of Agronomic traits in backcrossed population

간장의 경우 평균을 중심으로 2개의 굴곡이 보이며 수장과 수수는 C3G 함량 만큼은 아니지만 왼편으로 약간 치우친 모습을 보인다(그림 5.)

2) GBS를 이용한 genotyping 및 고밀도 지도 작성

GBS library를 제작하여 library QC를 측정한 결과(표 3) 평균 사이즈가 673bp로 NGS 분석을 위한 sequencing에 적합한 크기를 보였다.

Table 3. Result of NGS library QC

#	Library Name	Library Type	Conc. (ng/ul)	Conc. (nM)	Size (bp)
1	R1	GBS library(double enzyme)	28.02	64.06	673
2	R2	GBS library(double enzyme)	32.75	74.06	681

여교배 집단 225개 계통에서 기능성 성분 관련 표현형을 기준으로 정규 분포가 되도록 96 계통을 선발하여 기본 유전자지도 작성에 활용하였다. 분석결과 총 6,703개의 SNPs가 측정되었으며 그 위치는 그림 6과 같다.

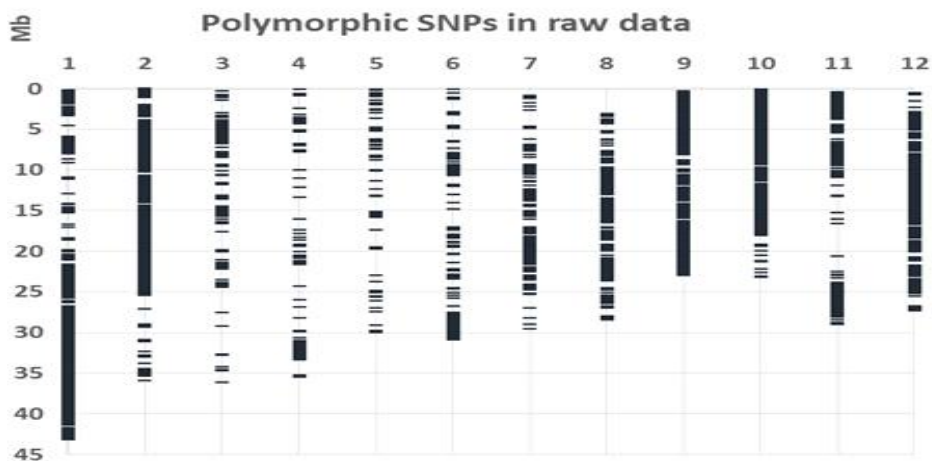


Fig. 6. Physical position fo 6,703 SNPs generated by GBS analysis

발굴된 총 6,703개의 SNPs 마커중 missing data 인 N의 개수가 평균 (49개) 이상인 SNP를 제거하여 3,737 SNPs 로 필터링하였고, 분석된 집단 96개의 line 중 missing data N이 3,737 개중 2,980개 이상인 24 lines를 제외하였다. 이후 2차 필터링으로 N이 22개 이상인 SNPs를 제외하여 3,269 SNPs를 부분형(paternal)은 2, 혼합형(hetero)은 1, 모본형(maternal)은 0 그리고 missing은 -1로 바꾸어 집단내에서 분리양상을 분석하였다. 이를 기초로 3차 필터링을 하여 2,951 SNPs를 제외하였고 redundant SNP binning을 거쳐 최종적으로 475 SNPs 와 72 lines을 이용하여 mapping과 QTL 분석을 진행하였다.

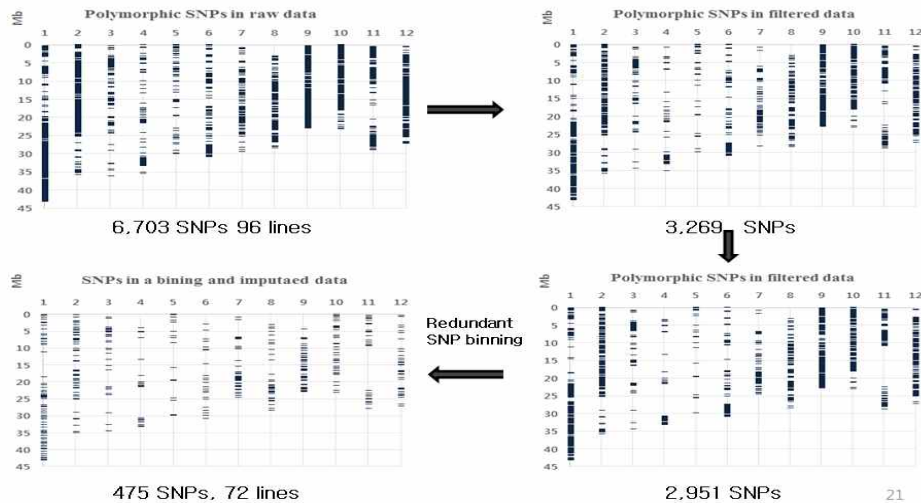


Fig. 7. Filtering and binning the GBS-SNPs

그림 7은 filtering 등을 거치면서 SNP의 물리적 위치를 표시한 그림이다. 의미있는 마커를 사용하기 위하여 많은 SNP와 계통들을 제거하였다.

최종적으로 분석집단을 genotyping 한 모습은 아래 그림과 같다(그림 8).

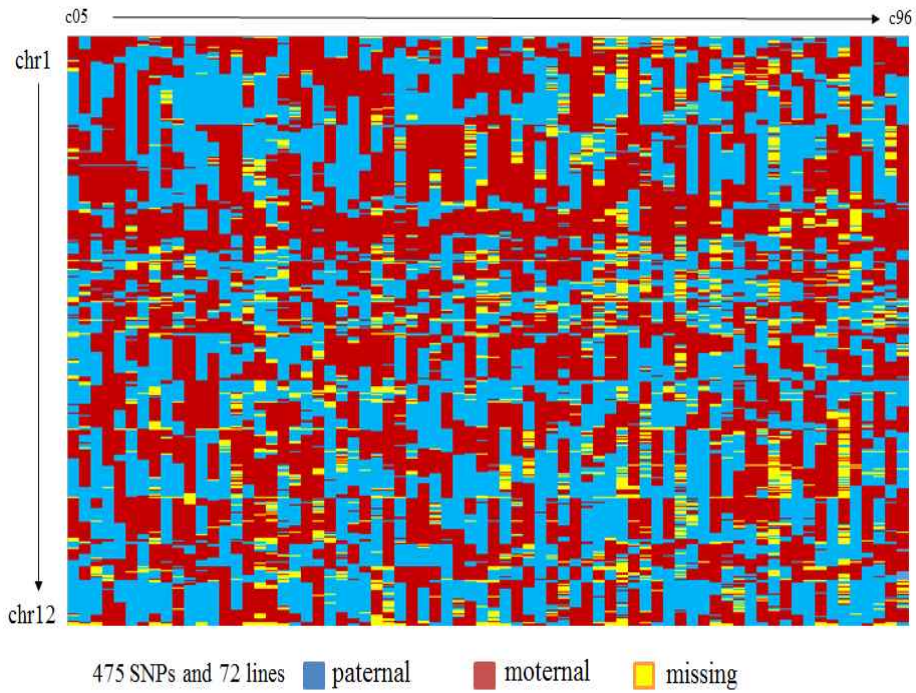


Fig. 8. Genotyping of the population after filtered and binned the SNPs

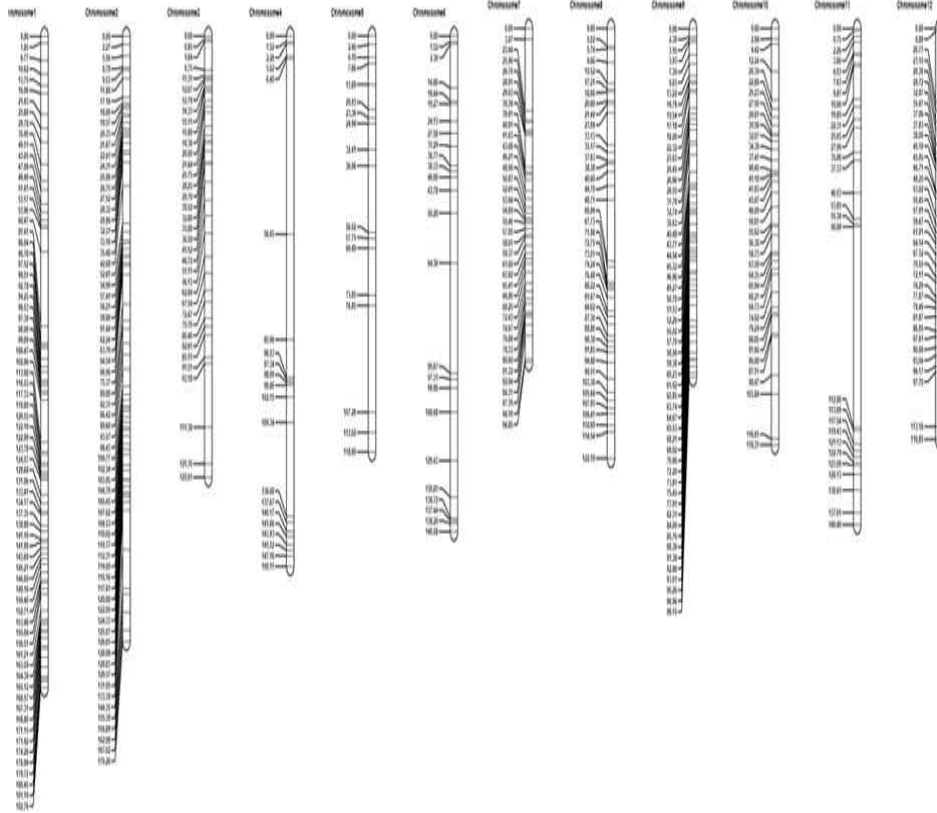


Fig. 9. Physical map position of 475 SNPs

475개의 SNP 위치 정보를 기초로 작성한 physical map은 그림 9와 같다. 4번 염색체와 11번 염색체의 중앙 부위에 gap 이 다소 크게 나타났는데 이는 사용된 제한효소 특이성 때문에 나타나는 결과로 생각된다(He J. et al. 2014). Furuta T. et al.(2016) 에서도 4번 염색체의 gap이 보고되었다.

Table 4. Linkage map on 12 chromosomes

Linkage group	No. of SNPs	Length	Ave. distance between SNPs	Linkage group	No. of SNPs	Length	Ave. distance between SNPs
Ch. 1-1	18	71.7	4.0	Ch. 6	24	149.0	6.2
Ch. 1-2	54	140.0	2.6	Ch. 7	38	128.9	3.4
Ch. 2	68	223.8	3.3	Ch. 8	40	138.3	3.5
Ch. 3	36	161.1	4.5	Ch. 9	55	145.0	2.6
Ch. 4-1	5	8.0	1.6	Ch. 10	37	140.4	3.8
Ch. 4-2	7	26.6	3.8	Ch.11-1	18	63.2	3.5
Ch. 4-3	8	14.9	1.9	Ch.11-2	11	31.9	2.9
Ch. 5-1	15	99.3	6.6	Ch. 12	36	126.8	3.5
Ch. 5-2	3	11.9	4.0				
Total					473	1680.8	3.55

SNP 마커의 유전자형 결과를 토대로 Joinmap 프로그램을 이용하여 linkage map을 작성하였다. 1번, 5번 11번 염색체는 2개 그룹으로 나뉘어졌으며 물리지도상에 gap 컷던 4번 염색체는 3개 그룹으로 나누어져 총 17개의 연관군을 보였다. Mapping된 마커는 473개이고, 마커간 평균 거리는 3.55cM 이었다.

3) 농업형질 및 C3G 함량 관련 QTL 분석

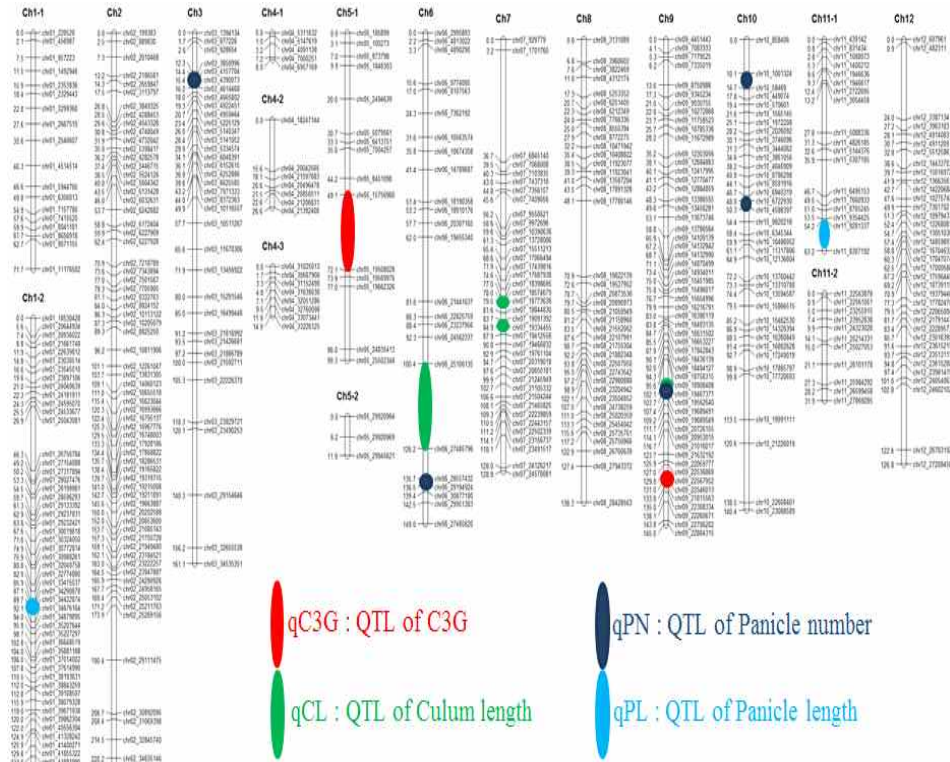


Fig. 10. Linkage map of 12 chromosomes and QTLs for 3 agronomic traits and C3G contents

총 17개의 linkage group의 map상 위치를 보면 일부 SNP의 위치가 physical map과 다르게 나타났는데(그림 10) 이는 분석 집단이 작고 여교배 집단이기 때문이라고 생각된다. 그러나 QTL 분석을 위한 지도로는 충분히 even space 하게 배치된 것으로 판단하여 QTL cartographer를

이용하여 QTL 분석을 수행하였다. QTL 분석 결과 간장 관련 QTL은 6번 염색체에 1개, 7번 염색체에 2개, 9번 염색체에 1개로 총 4군데에 위치하였다. 수장 관련 QTL은 1번 염색체와 11번 염색체에 1개씩 총 2개, 수수 관련 QTL은 3번, 6번, 9번 그리고 10번 염색체에 5개가 위치하였다. C3G 관련 QTL은 5번 염색체와 9번 염색체에 각 1개씩 총 2개가 탐색되어 총 13개의 QTL이 탐색되었다(표 5).

Table 5. QTLs of agronomic traits and C3G content

	Linkage	Position	LOD	Additive	R ²	Flanking marker
Culm length	6	101.89	3.90	-8.3326	0.1735	Chr06_25106135~ chr06_27485796
	7	77.06	3.44	6.9972	0.1460	Chr07_18398695~ chr07_18574679
	7	84.87	4.47	7.5143	0.1722	Chr07_19334455~ chr07_19412558
	9	103.67	4.10	7.7319	0.1737	Chr09_19562540~ chr09_19689491
Panicle length	1-2	87.09	3.00	-1.0688	0.1257	Chr01_34290878~ chr01_34422874
	11-1	53.47	3.50	1.1411	0.1447	Chr11_9354425~ chr11_9281337
Panicle number	3	12.30	3.92	-1.8468	0.1361	Chr03_3850996~ chr03_4157704
	6	136.55	4.16	1.7612	0.1509	Chr06_29194924~ chr06_30877180
	9	102.07	3.43	1.6249	0.1238	Chr09_18908408~ chr09_19467371
	10	11.14	3.41	2.3011	0.1340	Chr10_1001324~ chr10_58469
	10	40.83	5.66	-3.0438	0.2146	Chr10_8786298~ chr10_8591916
C3G content	5-1	54.6	3.99	-112.9	0.2344	Chr05_15756960~ chr05_19508028
	9	128.0	3.61	93.0	0.1583	Chr09_22567952~ chr09_22546013

C3G 관련 QTL은 5번과 9번에서 탐색되었다. 5번 QTL은 전체 변이의 23% 를 설명하는 QTL로서 수원 425의 대립유전자가 C3G 함량은 증가시켰고, 9번 염색체는 전체 변이의 15.8%를 설명하는 것으로 C3GHi 즉 흑진주에서 유래된 QTL 로 분석되었다.

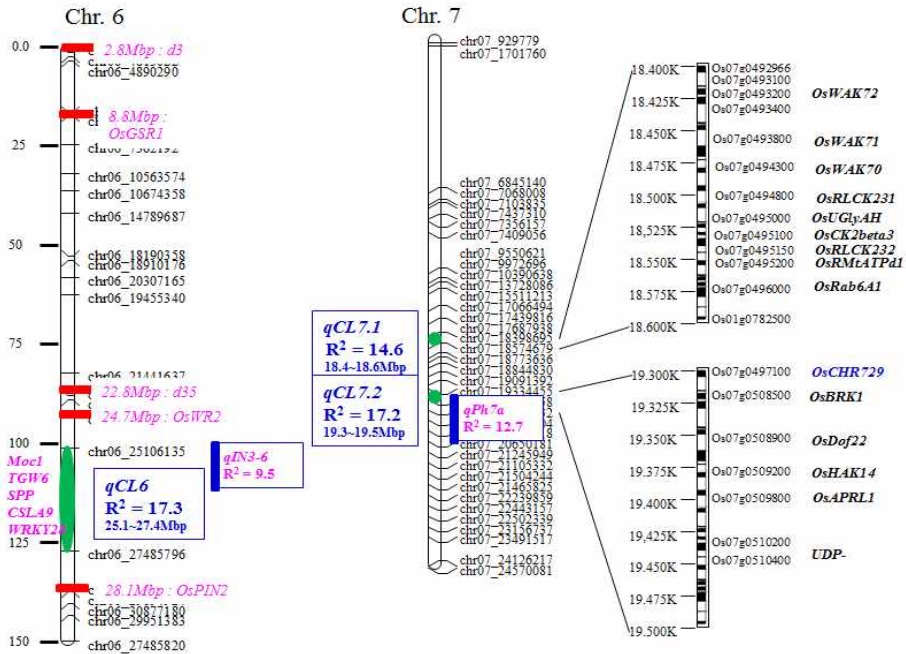


Fig. 11. QTLs of culm length on Chr. 6 and Chr. 7

간장에 대한 QTL 분석 결과 6번 염색체 하단에서 탐색되었다(그림 11). 6번 염색체에서는 dwarf 와 관련된 *d3*, *d35* 등이 보고되었고, 24.7 mbp 근처에서는 *wax synthesis regulator* 등이 보고되었다. QTL 영역내 보고된 유전자는 *MOC1*(분얼수), *TGW6*(천립중), *SPP*(뿌리 신장) 을 비롯한 *CSLA9* 와 *WRKY28* 등 초장과 관련된 몇 개 유전자가 보고된 바 있다. 또한 2개의 병저항성에 관련된 *LRR*이 같은 위치에서 보고되었다. 또한 고시히까리 / kasarath BIL에서 보고된 *qIn3* (3번째 절간 길이 관련) QTL 이 보고된 바 있다.

7번 염색체에서는 2군데 QTL이 탐색되었다. 7.1 부위에는 세포벽 관련

유전자(*WAK*)와 *RLCK* (알루미늄 내성 관련 유전자) 등이 존재하였다. 7.2 부위에는 클로로프라스트 발달과 관련된 *CHR729* 가 존재하였는데 식물체의 크기와 관련된 것으로 알려져 있다. 이 부위에는 plant height와 관련된 QTL도 보고된 바 있다.

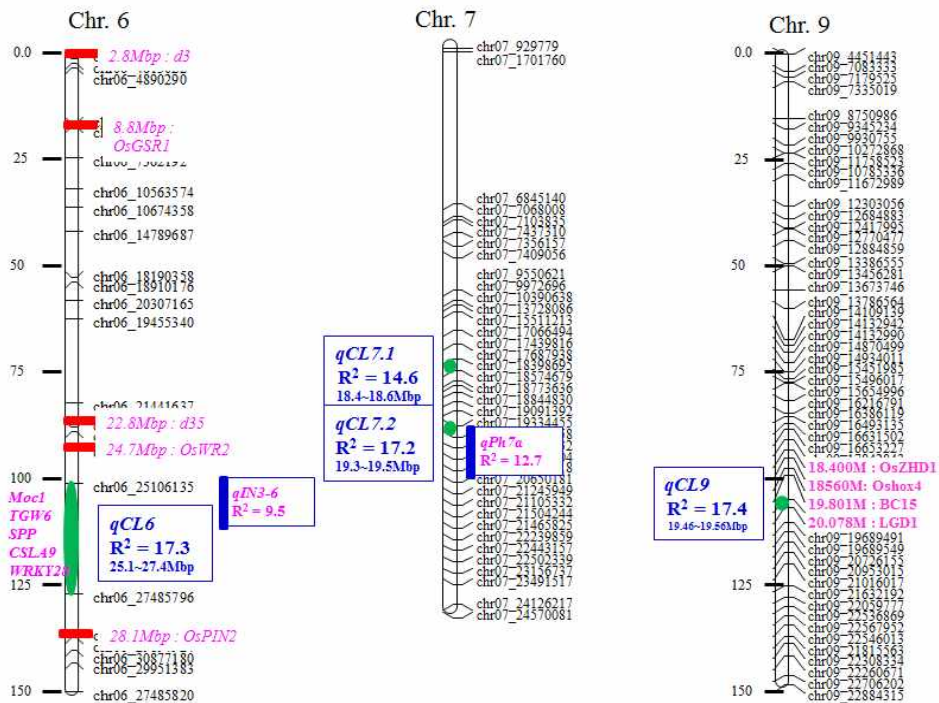


Fig 12. QTLs of Culm length on Chr. 6, 7 and 9

9번 염색체 아랫 부분에서 탐색된 QTL 은 R^2 가 17.4 정도였다(그림 12). 동 위치에는 dwarf 와 관련된 *OsZHD1* 과 *Os hox4* 가 보고된 바 있고, *brittle culm(BC)* 유전자가 위치하고 있다. 그리고 분얼수와 이삭길이에 영향을 주는 것으로 알려진 *LGD1* 유전자가 보고된 바 있다.

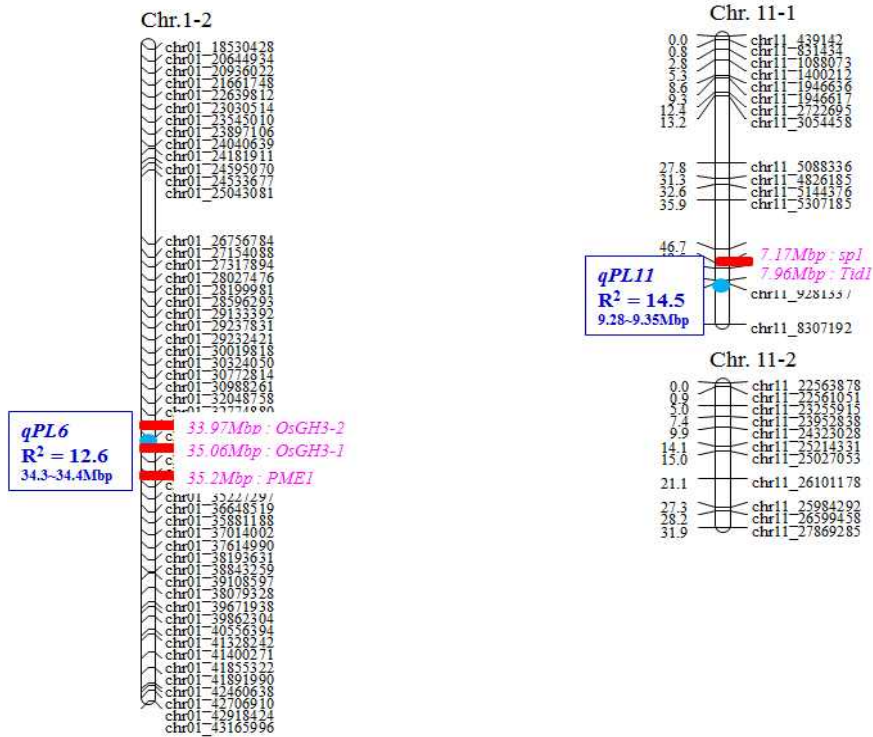


Fig. 13. QTLs of panicle length on Chr. 1 and Chr. 11

이삭 길이 관련 해서는 2개의 QTL 이 탐색되었다(그림 13). 1번 염색체 하단에서 12.6 % 의 QTL 이 11번 염색체 중단에서 14.5% 의 QTL이 탐색 되었다.

1번 염색체 QTL 부위에는 단간 및 분얼과 관련된 *OsGH* 와 잎 노화와 관련된 *PME1* 유전자가 보고된 바 있으나, QTL은 보고된 바 없다.

11번 염색체의 중단은 이삭 길이 감소와 관련된 *sp1* 과 *Tid1* 유전자가 보고된 자리이다.

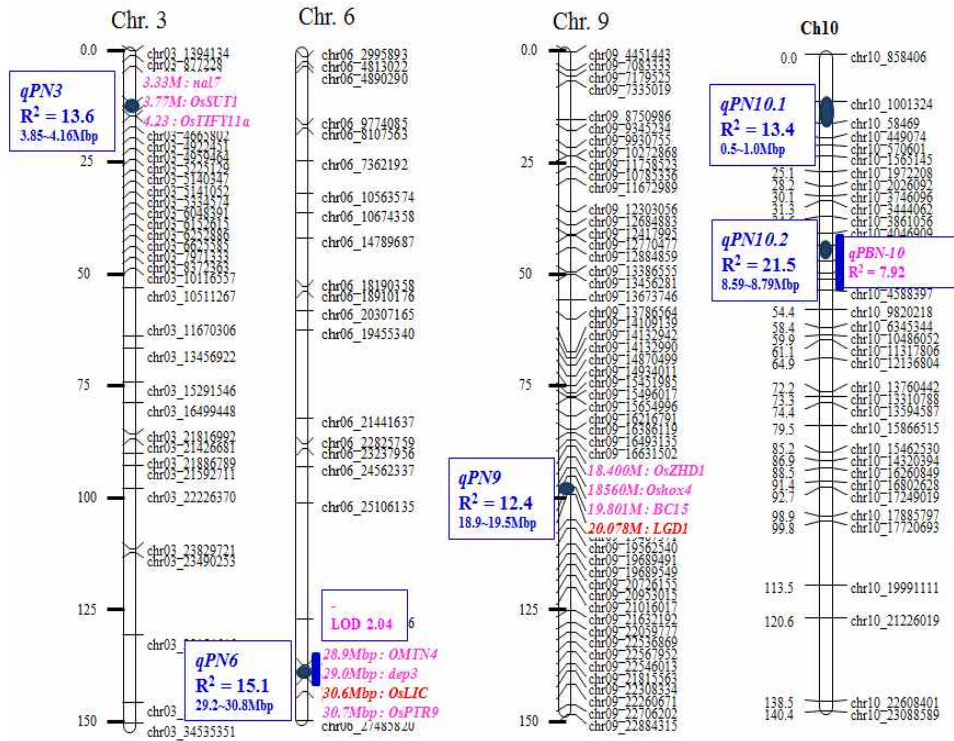


Fig. 14. QTLs of panicle number on Chr. 3, 6, 9 and 10

분얼 수 관련된 QTL은 총 5개로 3번, 6번, 9번 염색체에 1개씩 위치했으며, 10번 염색체에 2개의 QTL이 위치하였다(그림 14).

3번 염색체 상단은 13.6% 정도 변이를 포함하는 영역으로 주로 단간, 출수 관련된 유전자 및 QTL이 보고된 자리이다.

6번 염색체 말단은 15.1% 의 변이를 포함하는 영역으로 분얼과 관련된 *OSLIC* 와 종자 및 이삭 크기와 관련 된 *OsPTR9* 이 보고된 영역이며, 또한 밀양 46호에서 이삭수 관련 QTL이 보고된 바 있다.

9번 염색체 부위는 간장에서 QTL이 보인 영역으로 분얼수와 이삭길이에

영향을 주는 것으로 알려진 *LGD1* 유전자가 보고된 바 있다.

10번 염색체에서는 2개 영역에서 각각 13.4, 21.5% 의 QTL이 탐색되었고, 10.2 영역에서는 1차 분얼수와 관련된 QTL이 보고된 바 있다.

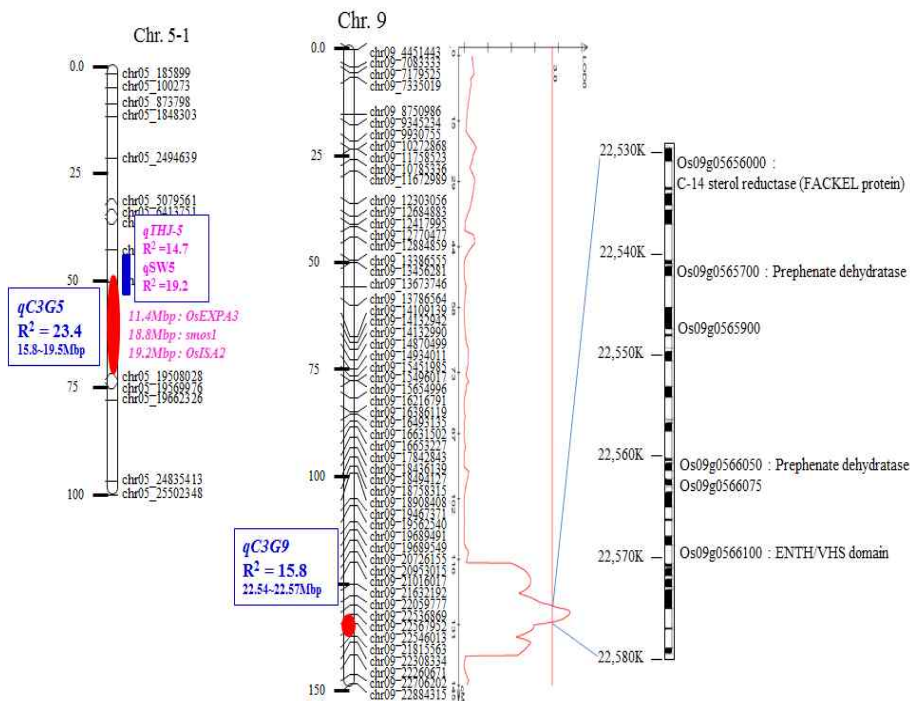


Fig. 15. QTLs of cyanidin 3-glucoside content on Chr. 5 and Chr. 9.

C3G와 관련된 QTL은 5번과 9번 염색체에서 탐색되었고, 각각의 R^2 값은 23.4, 15.8로 나타났다(그림 15).

5번 염색체에서 탐색된 qC3G5는 수원 425호 대립유전자가 C3G 함량을

증가시킨 것이고, 9번 염색체 부위는 C3GHi의 대립유전자, 즉 흑진주의 대립유전자가 C3G 함량을 높이는 것으로 나타났다.

5번 염색체 부위에는 종자 두께와 관련된 QTL이 2개 보고 되었고, 세포벽 관련 *Os EXPA3* 유전자와 세포크기를 작게하여 세포수를 증가시키는 *smos1* 유전자가 보고된 바 있다.

9번 염색체 말단 부위에서 탐색된 QTL에는 C-14 *sterol reductase* 등이 보고되었다.

4) Single Effect ANOVA 분석을 통한 연관부위 탐색.

475개 SNP 마커와 72 계통의 연관성분석을 위하여 Single Effect ANOVA로 분석한 결과 간장형질과 관련하여 염색체 2번, 6번 그리고 7번에서 3개의 연관된 위치가 탐색되었다. 수장은 1번, 4번, 11번 그리고 12번에서 연관된 부위가 탐색되었으며, 수수형질과 관련해서는 3번에 2개, 5번, 6번 그리고 11번에 1개씩 관련 위치가 탐색되었다. C3G 함량 관련해서는 2번 5번, 8번, 9번, 10번, 12번에 1개씩 5부위가 탐색되었다.

Table 6. QTLs of Culm length analyzed by single effect ANOVA

QTLs	Chr.	bp	cM	F	Prob	Sig	Rsq	Add
qcl2	2	19,215,008	110.0	6.24	0.015	*	0.085	5.2
	2	19,319,715	110.8	5.60	0.021	*	0.078	4.9
	2	19,663,887	112.3	4.76	0.033	*	0.067	4.6
	2	20,202,588	114.7	6.74	0.015	*	0.090	5.7
	2	20,653,600	116.2	4.32	0.041	*	0.059	4.6
qcl6	6	23,237,956	97.3	8.34	0.005	**	0.119	5.9
	6	24,562,337	99.9	4.03	0.049	*	0.058	4.3
	6	25,106,135	106.7	4.36	0.041	*	0.062	4.6
qcl7	7	17,687,938	52.4	7.63	0.007	**	0.104	-5.8
	7	18,398,695	53.9	11.48	0.001	**	0.148	-6.7
	7	18,574,679	54.7	8.63	0.005	**	0.116	-5.9
	7	18,773,636	55.5	10.79	0.002	**	0.140	-6.4
	7	18,844,830	57.1	8.40	0.005	**	0.114	-5.7
	7	19,091,392	58.6	10.46	0.002	**	0.135	-6.2
	7	19,334,455	59.4	11.85	0.001	***	0.150	-6.7
	7	19,412,558	61.7	9.28	0.003	**	0.120	-6.1
	7	19,466,832	63.9	6.06	0.016	*	0.082	-5.0
	7	19,761,104	65.4	8.91	0.004	**	0.114	-5.8
	7	20,319,018	66.9	8.44	0.005	**	0.112	-5.7
	7	20,650,181	69.3	8.42	0.005	**	0.113	-5.9
	7	21,105,332	72.4	3.78	0.056	#	0.052	-4.0
	7	21,245,949	74.7	7.21	0.001	**	0.095	-5.3

간장길 이와 관련해서는 2번, 6번 7번 염색체에서 총 3개의 연관된 부위가 탐색되었다(표 6). QTL cartographer 로 QTL 분석하였을 때는 6번, 7번, 9번 염색체에서 7번에 2개로 총 4개의 QTL 위치가 탐색되었으나(표

3), Single effect ANOVA 분석으로는 2번에 염색체에 새로운 QTL 한개가 탐색이 되었고 7번은 2개에서 1개로 분석이 되었다. 9번 염색체에 있던 QTL은 유의하지 않았다. 위치가 겹쳐서 나타난 6번과 7번의 위치는 약간의 차이만 있을 뿐 대체적으로는 비슷하였다.

Table 7. QTLs of Panicle length analyzed by single effect ANOVA

QTLs	Chr.	bp	cM	F	Prob	Sig	Rsq	Add
qpl1	1	33,415,037	141.2	8.86	0.004	**	0.115	1.0
	1	34,290,878	142.0	7.49	0.008	**	0.108	1.0
qpl4	4	20,042,565	96.5	6.13	0.016	*	0.091	0.8
	4	20,496,478	97.3	4.10	0.047	*	0.058	0.7
	4	20,850,511	98.1	5.59	0.021	*	0.079	0.5
	4	21,206,831	98.9	4.90	0.030	*	0.070	0.6
qpl11	11	8,765,245	53.9	7.12	0.001	**	0.096	-0.9
	11	9,281,337	55.3	8.61	0.005	**	0.112	-1.0
	11	9,354,425	56.1	8.75	0.004	**	0.114	-1.0
qpl12	12	26,703,192	113.2	5.14	0.003	*	0.093	1.0
	12	27,208,450	116.9	4.38	0.004	*	0.087	0.9

수장의 경우는 1번, 4번, 11번, 12번 염색체 위에 4개의 연관된 부위가 탐색되었다(표 7). 앞선 분석에서 2군데 위치만 탐색되었으나 연관분석으로 4번과 12번에 1개씩 QTL이 추가로 탐색되었다.

Table 8. QTLs of Panicle number analyzed by single effect ANOVA

QTLs	Chr.	bp	cM	F	Prob	Sig	Rsq	Add
qpn3.1	3	3,850,996	9.8	7.36	0.008	**	0.102	1.6
	3	4,157,704	11.3	8.26	0.005	**	0.113	1.7
	3	4,390,073	12.1	7.89	0.006	**	0.104	1.7
qpn3.2	3	23,490,253	91.5	4.46	0.039	*	0.069	-1.1
	3	23,829,721	93.2	5.01	0.029	*	0.070	-1.1
qpn5	5	100,273	0.0	9.16	0.004	**	0.125	-1.5
	5	185,899	2.5	3.31	0.074	#	0.048	-0.9
	5	873,798	6.1	3.84	0.055	#	0.061	-1.0
	5	1,848,303	7.9	6.22	0.015	*	0.091	-1.2
	5	2,494,639	13.9	4.97	0.030	*	0.083	-1.2
qpn6	6	27,485,820	130.8	6.91	0.011	*	0.112	-1.5
	6	28,557,432	136.7	5.96	0.017	*	0.080	-1.2
	6	29,194,924	137.4	6.97	0.010	*	0.093	-1.4
	6	29,951,383	138.2	6.79	0.011	*	0.092	-1.3
qpn11	11	9,281,337	55.3	4.42	0.039	*	0.061	1.0
	11	9,354,425	56.1	8.18	0.006	**	0.107	1.4

수수 형질의 single effect ANOVA 분석 결과 9번과 10번에 위치했었던 QTL은 탐색이 되지 않았고, 3번 염색체에 추가 1개와 5번, 11번 염색체에 수수 관련 QTL이 탐색이 되어 총 5개의 QTL이 탐색되었다(표 8).

Table 9. QTLs of C3G content analyzed by single effect ANOVA

QTLs	Chr.	bp	cM	F	Prob	Sig	Rsq	Add
qc3g2	2	10,655,518	67.0	0.85	0.361		0.012	29.9
	2	10,811,906	75.4	7.81	0.007	**	0.102	86.6
	2	12,261,047	80.0	4.29	0.042	*	0.058	66.4
qc3g5	5	24,835,413	73.9	5.74	0.020	*	0.085	81.0
	5	25,502,348	76.9	5.16	0.027	*	0.090	96.1
qc3g8	8	3,822,469	5.0	5.00	0.029	*	0.068	-75.8
	8	3,960,602	5.7	4.91	0.030	*	0.067	-76.2
qc3g9	9	22,567,952	95.2	6.99	0.010	*	0.100	-92.6
	9	22,706,202	97.0	7.71	0.007	**	0.117	-98.7
	9	22,884,315	98.2	5.01	0.031	*	0.106	-95.9
qc3g10	10	17,720,693	86.7	4.49	0.038	*	0.063	68.3
	10	19,991,111	98.5	5.98	0.017	*	0.089	84.3
	10	21,226,019	103.8	6.29	0.015	*	0.088	81.7
qc3g12	12	4,914,083	28.7	3.10	0.083	#	0.045	58.3
	12	5,512,586	33.0	5.94	0.018	*	0.087	92.0
	12	7,301,702	35.5	3.35	0.072	#	0.049	67.0

C3G 함량과 관련하여 2번, 5번, 8번, 9번, 10번 그리고 12번 염색체에서 연관된 부위가 탐색되었다. 12번 염색체에서는 2곳으로 탐색되었다. 8번, 9번 염색체 부위는 모본 즉 C3GHi 대립유전자 효과이고 2, 5, 10, 12번

염색체는 부분 즉 수원 425호 대립유전자 효과로 분석되었다(표 9). 2번과 9번 위치는 R^2 값이 고도로 유의하고 additive 한 효과도 크게 나타났다.

C3G 함량과 연관된 QTL을 1개 이상 보유한 계통의 C3G 함량을 비교하였다(그림 16). 6개 QTL을 모두 가진 계통에서 C3G 함량이 높게 나타났으며 대체적으로 C3G 함량과 연관된 QTL을 많이 가진 개체의 C3G 함량이 높게 나타났다. 그러나 C10과 C41도 동일하게 C3G함량 연관 QTL 6개를 모두 갖고 있으나 C3G 함량에 차이가 나는 것으로 보아 QTL 분석으로 탐색 되지 않은 minor QTL 등이 영향을 준 것으로 생각된다.

Line No	C3G content	qc3g2	qc3g5	qc3g8	qc3g9	qc3g10	qc3g12
C3GHi	767.54						
S. 425	74.40						
C10	6	1,711.79					
C41	6	1,151.32					
C35	5	791.93					
C56	4	683.06					
C83	4	617.88					
C72	4	585.42					
C69	5	530.36					
C29	2	458.85					
C79	4	456.38					
C71	3	329.07					
C31	3	328.98					
C95	1	302.25					
C80	3	279.37					
C68	2	273.71					
C57	1	258.69					
C22	2	247.65					
C08	2	240.47					
C21	2	224.47					
C30	3	214.09					
C54	2	179.26					
C90	4	168.71					
C92	3	155.44					
C42	1	113.26					
C24	1	110.70					
C67	1	102.45					
C59	1	46.32					
C60	1	30.72					

Fig. 16. Comparison of C3G content in lines which has more than 1 QTLs

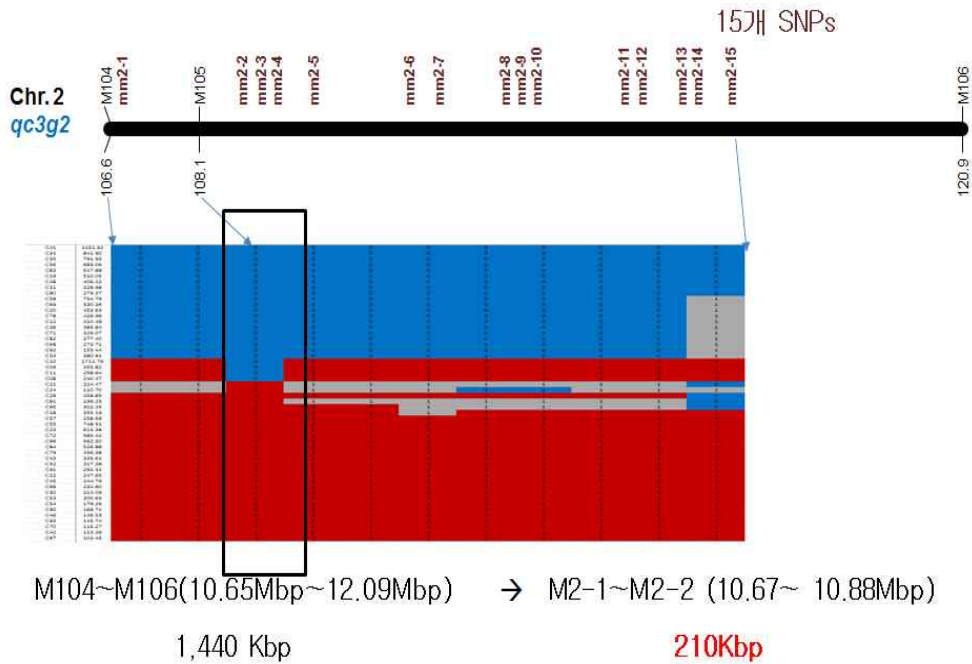


Fig. 17. Narrow down process in QTL qc3g2 region using eliminated SNPs during filtering.

C3G 함량과 연관된 QTL의 범위가 넓게 분석되어 관련된 후보 유전자를 찾기 위하여 각 QTL 범위를 좁히는 작업을 진행하였다. GBS 분석 결과와 GBS 분석에 이용되지 않았던 추가 후대 계통의 유전자 정보를 활용하여 각각의 QTL 영역을 좁혔다(그림 17). 필터링 등을 통해 제외된 SNP 마커들을 이용하여 발굴된 QTL내에 교차가 일어난 계통들을 대상으로 거리를

좁히는 마커로 사용하였다.

5번과 8번 염색체에 위치한 C3G 함량 관련 QTL 범위도 같은 방법으로 줄여 나갔다(그림 18과 그림 19). 5번 염색체에 위치한 QTL 범위는 10Mbp 에서 7kbp로 범위를 줄였으며, 8번 염색체에 위치한 QTL 범위는 2Mbp에서 2.2kbp로 범위를 줄였다. 9번과 10번 12번 염색체의 QTL도 범위를 줄여나갔다(그림 20, 21, 22).

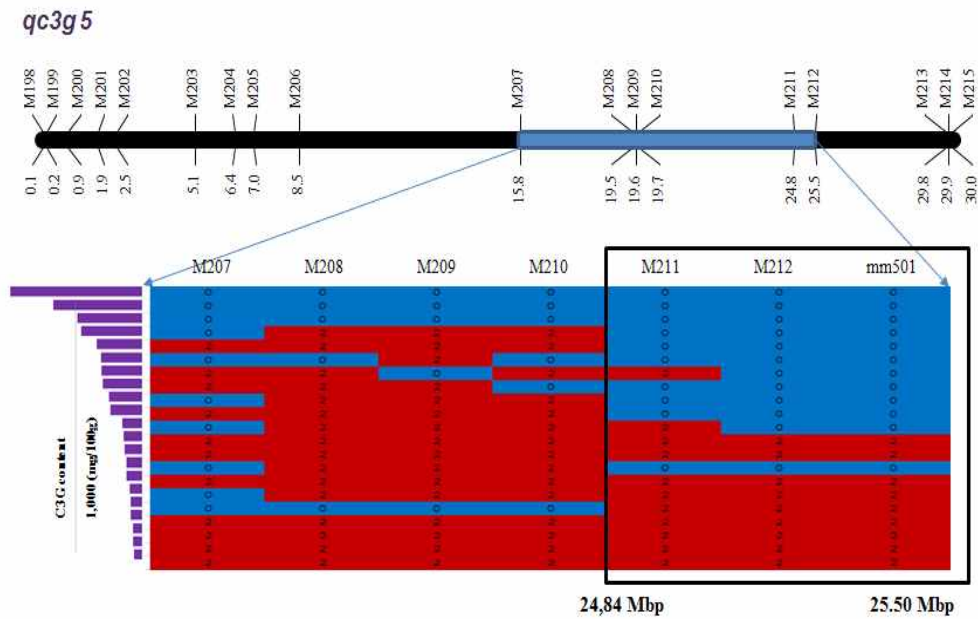


Fig. 18. Narrow down process in *qc3g5* QTL region on chromosome 5.

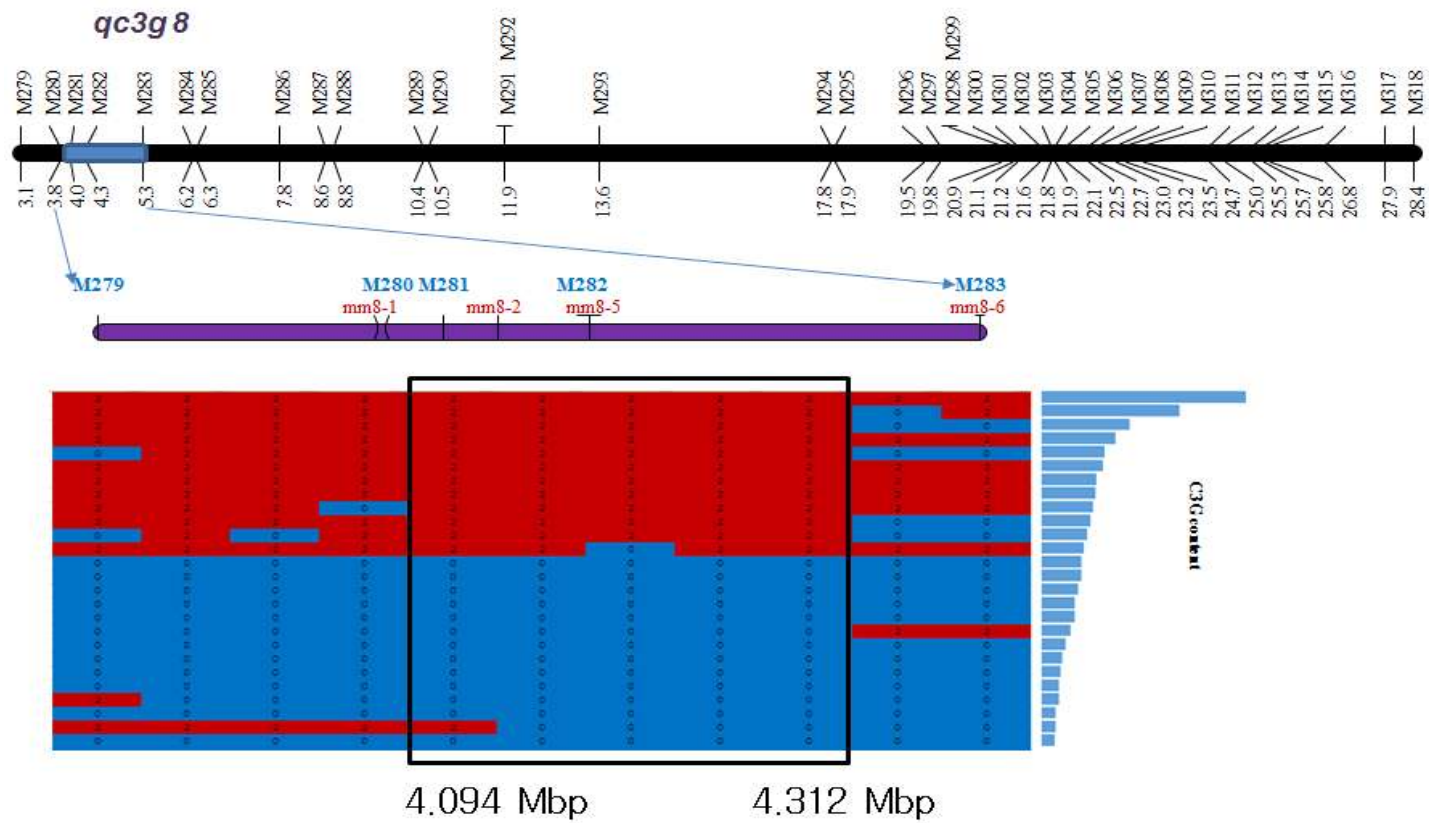


Fig. 19. Narrow down process in qc3g8 QTL region on chromosom 8.

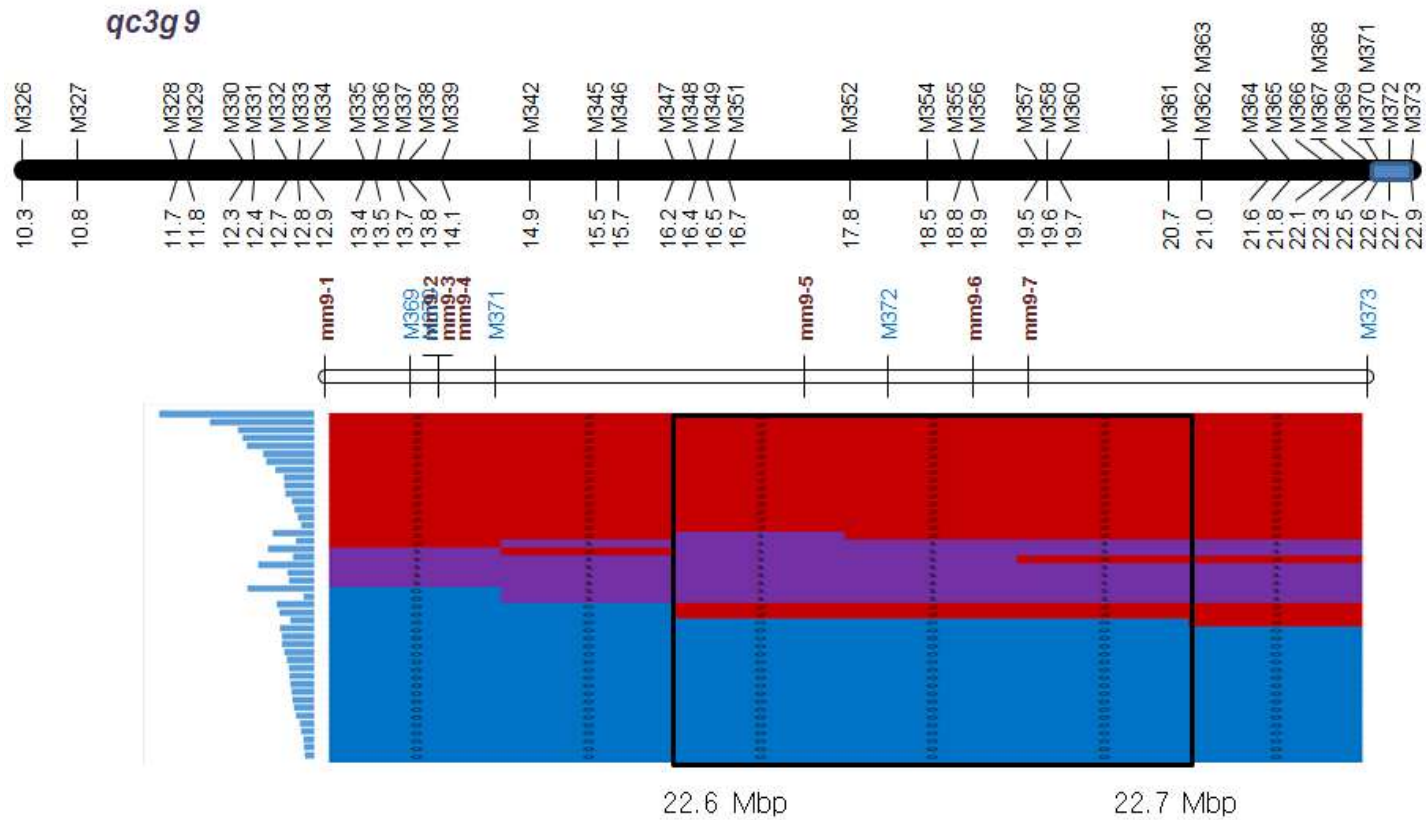


Fig. 20. Narrow down process in *qc3g9* QTL region on chromosome 9.

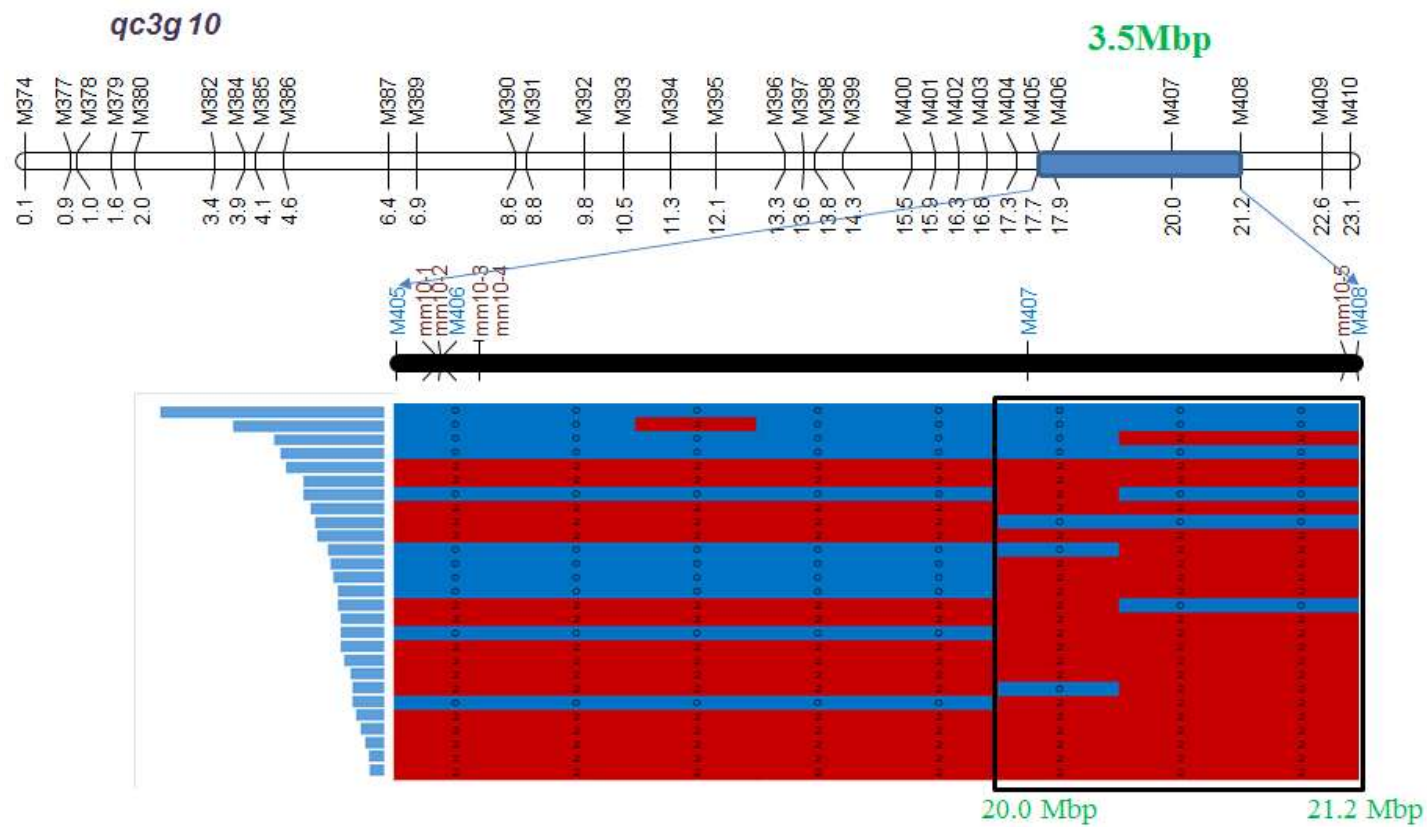


Fig. 21. Narrow down process in qc3g10 QTL region on chromosom 10.

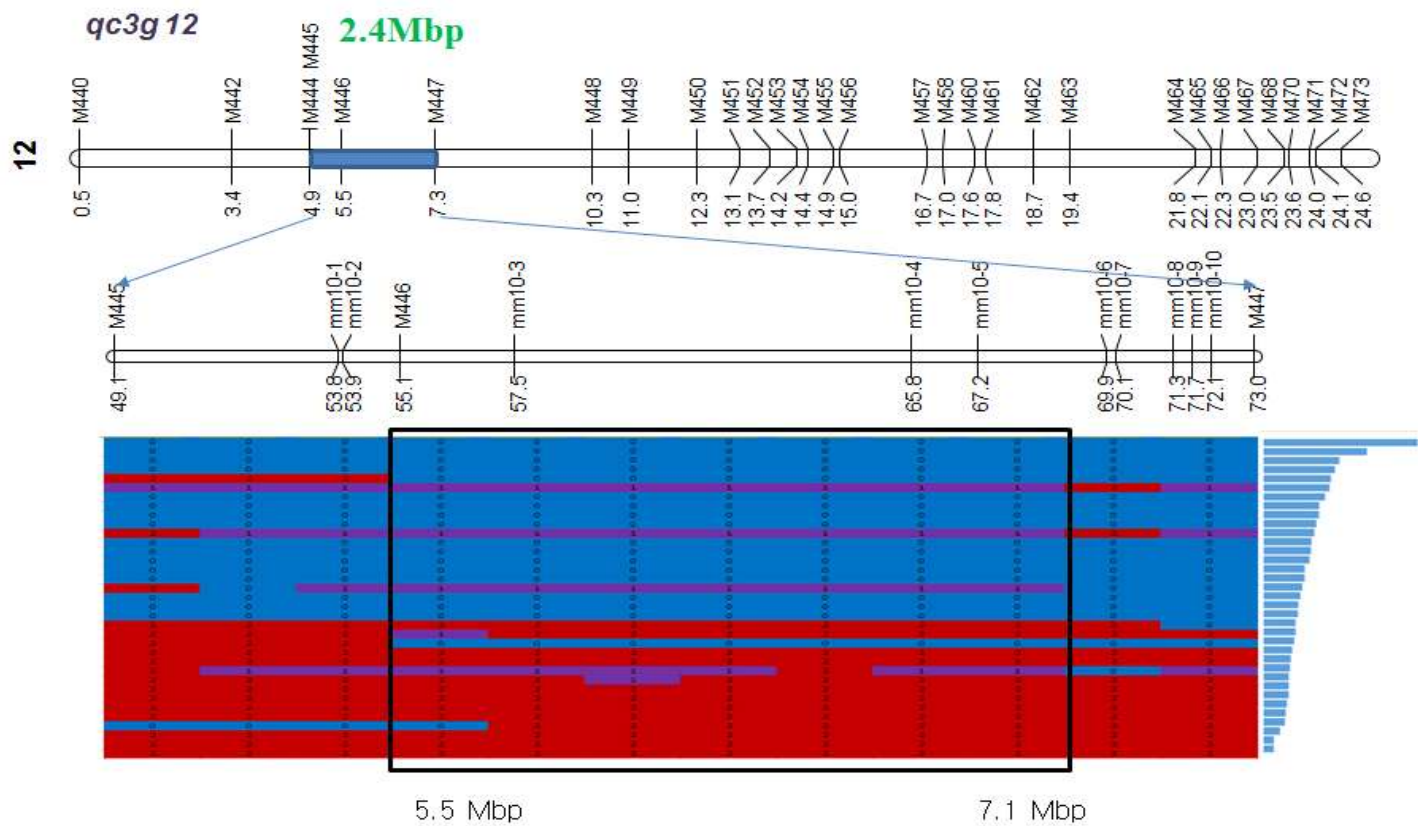


Fig. 22. Narrow down process in *qc3g12* QTL region.

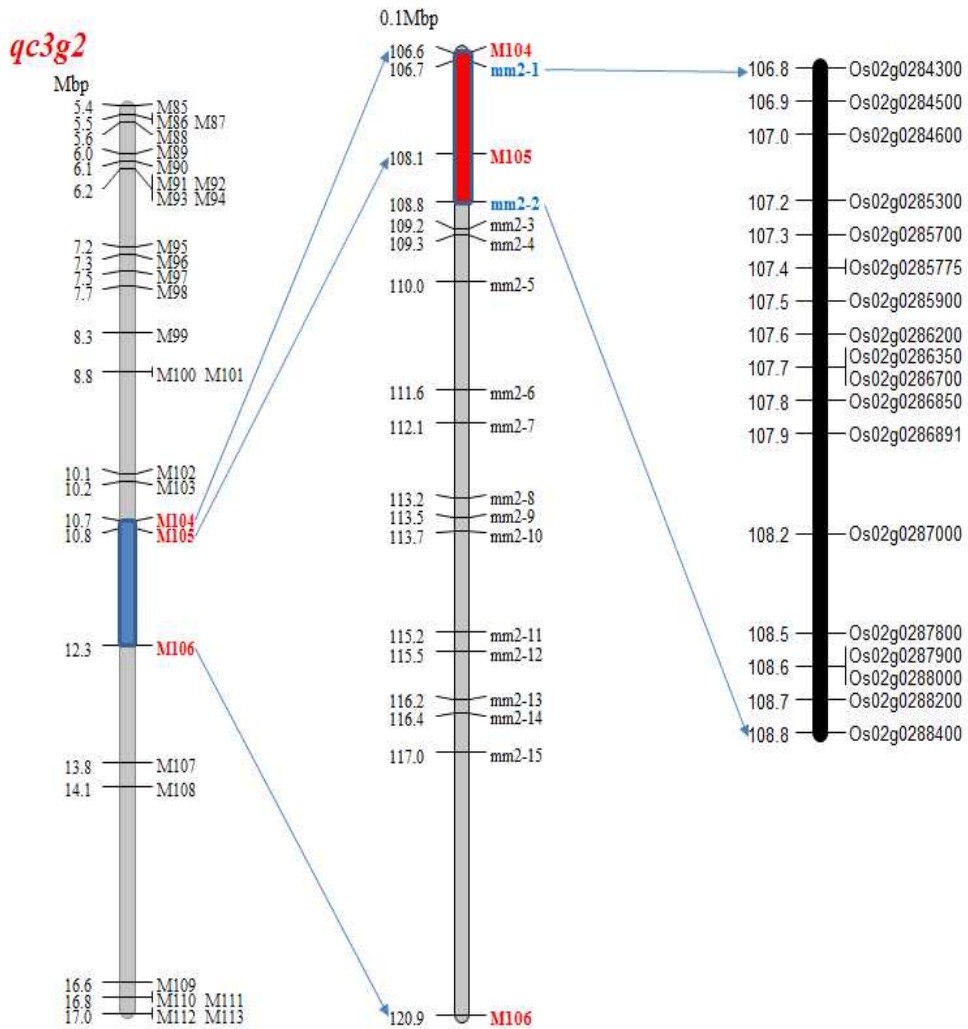


Fig. 23. Putative genes in *qc3g2* QTL region on chromosom 2.

R^2 값이 큰 2번과 9번 염색체에 위치한 QTL에서 대해서는 후보 유전자를 탐색해 보았다(부록 1). GBS 결과로 좁힌 영역 내에 여전히 다수의 유전자가 존재하고 있다(그림 23). 모본 C3GHi와 수원 425의 Re-seq. 결과

를 rice snp miner 프로그램을 사용해서 후보 영역내 대량의 snp 를 확보하고, 제한효소를 이용할 수 있는 다수의 CAPS 마커를 추가시켰다(그림 24와 그림 25).

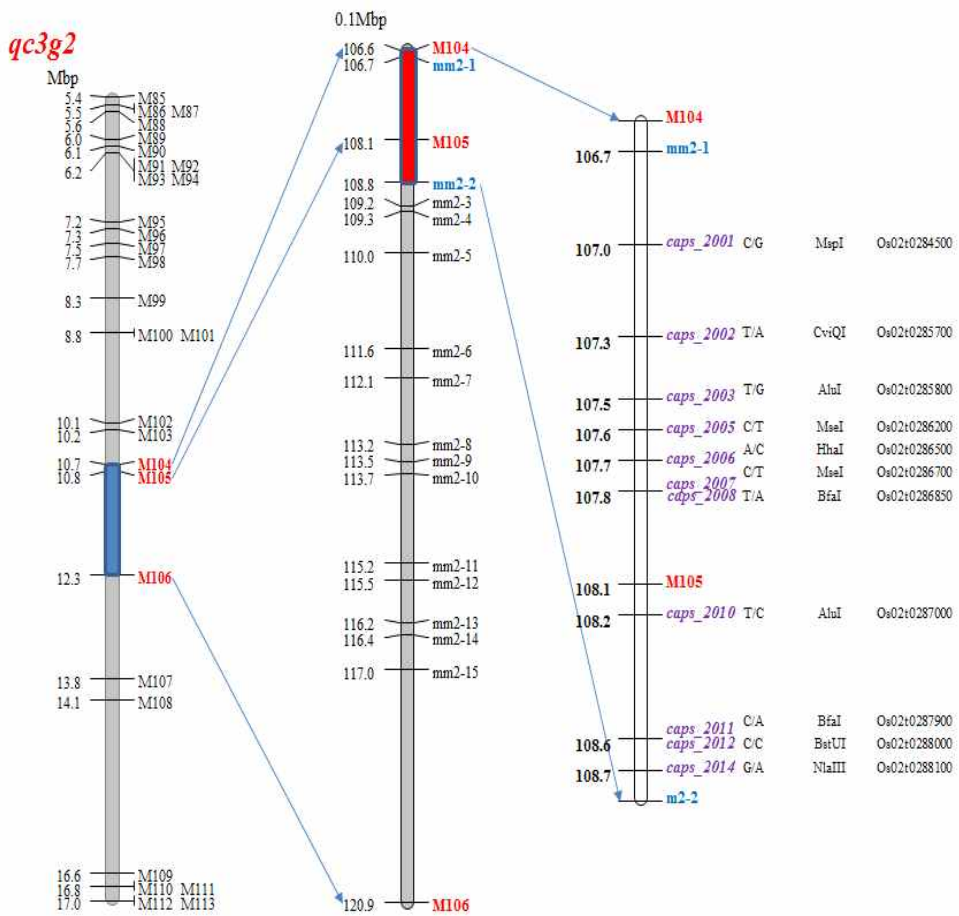


Fig. 24. CAPS markers in *qc3g2* QTL region on chromosome 2 using resequence data of parent C3GHi and S. 425.

모본에서 다형성평가를 수행한 CAPS 마커를 이용하여 목적영역 내에 교차지점이 있는지 여부를 살펴보았다. 앞에서 유전자형 분석에서 목적영역 내에서 교차가 있었던 계통을 대상으로 유전자형 평가를 통해 QTL내 범위를 좁혔다.

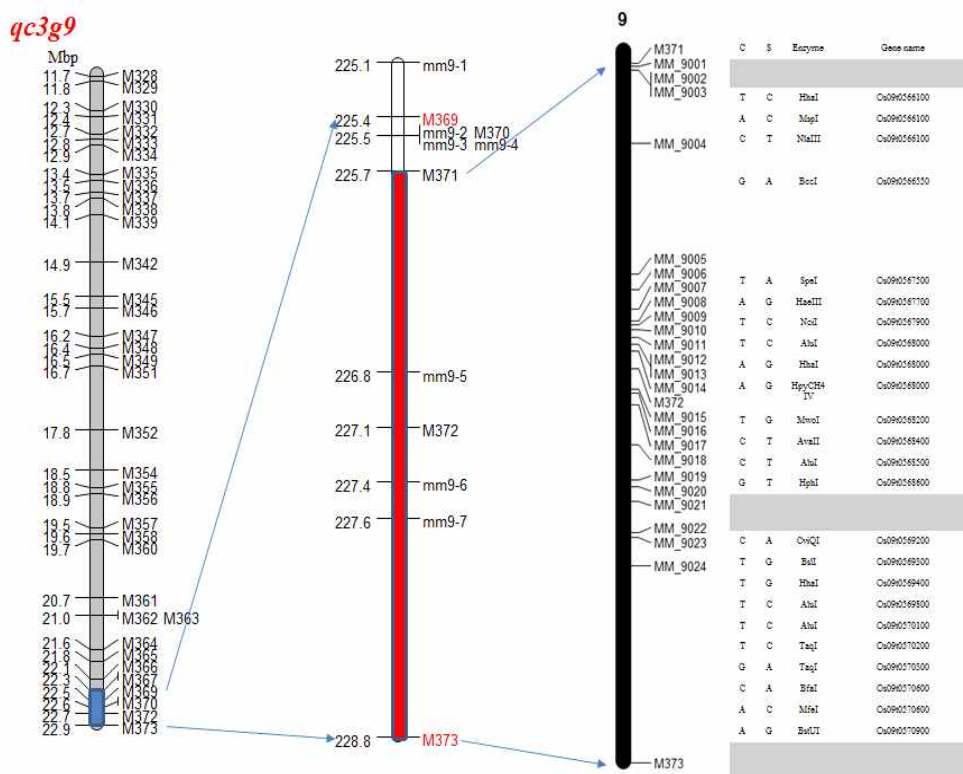


Fig. 25. CAPS markers in qc3g9 QTL region on chromosome 9 using resequence data of parent C3GHi and S. 425.

9번 염색체에서도 CAPS 마커를 개발하여 적용하였다.

5) C3G 함량 관련 후보유전자 haplotype 분석

Chr.2 OS02G0289500



Fig. 26. Haplotype analysis of 150 rice varieties to putative genes in qc3g2 QTL region.

QTL 영역을 좁히는 것과 별개로 후보 영역내 candidate 유전자에 대해 haplotype 분석을 수행하였다. 2번 염색체의 후보 유전자에 대해 국내육성 150품종에 대해 haplotype 분석을 수행한 결과이다. 크게 5개로 구분되었고, 자포니카는 1, 2 그룹에 인디카는 2, 3, 4, 5 그룹에 흩어져 나타났다(그림 26).

Chr. 9 OS09G0566100



Chr. 9 OS09G0566550

	0s09t0566550-01				
Position	22,603,029	22,604,233	22,604,894	22,605,463	22,607,304
Reference	T	G	G	T	C
genotype	C: 52 T: 101	A:21 G:132	A:21 G:132	A:21 T:132	0:132 1:21
Haplotype1 (통일형)	21개 (CAAA0)	한아름, 다산, 밀양 23호, 삼강, 통일, 등			
Haplotype2 (자포니카)	10 (CGGT0)	상남발, 농림나1호, 드래산, 녹양			
Haplotype3 (자포니카)	21 (CGGT1)	운봉, 신운봉, 중화, 그루, 금오, 진미, 적진주 등			
Haplotype4 (자포니카)	101 (TGGT0)	고품, 운광, 하이아미, 호품, 질보, 일품, 추정, 주남, 신통진 등			

Fig. 27. Haplotype analysis of 150 rice varieties to putative genes in qc3g9 QTL region.

9번 염색체의 후보 유전자에 대한 haplotype 분석 결과 대체적으로 대부분 유전자의 haplotype 은 생태형 즉, 자포니카와 통일형으로 구분되었고, 일부는 출수기와 관련성이 높아 보였다(그림 27). 본 실험의 교배조합이 인디카 극조생인 흑진주와 자포니카 중만생인 수원 425 조합의 결과로

생각되어지며 보다 구체적인 유전자의 기능에 대해서는 BIL 등을 육성할 필요가 있을 것으로 판단된다.

2 C3G 고함유 기능성 품종 육성

1) 출수기 및 주요 작물학적 특성

‘슈퍼자미2호’ 품종의 출수기는 중부평야 보통기 재배에서 8월 30일로 ‘슈퍼자미’ 품종보다 2일 늦은 만생종이다. 또한 ‘슈퍼자미2호’ 품종의 간장은 106.0 cm로 ‘흑진주벼’ 보다 28 cm가량 크며, ‘슈퍼자미’ 보다 약 30 cm 큰 품종이다. 수장은 20.4 cm로 ‘흑진주벼’ 보다 짧다. 포기당 이삭수는 8.5개로 ‘흑진주벼’와 비슷하고, 이삭당 벼알 수는 143.0개로 ‘흑진주벼’ 보다 많다. ‘슈퍼자미2호’의 임실율은 82.0% 수준이고, 현미 천립중은 30.1 g으로 ‘흑진주’보다 1.3배 가량 무거운 편이다(표 10).

Table 10. Major agronomic traits and yield components of rice variety Superjami 2

Cultivar	Year	Heading date	Culm Length (cm)	Panicle length (cm)	No. of panicle per hill	No. of grain per panicles	Ratio of fertility (%)	Brown rice TGW(g)
Heugjinju	2012	July.25	75.0	21.0	9.0	93.8	77.6	17.0
	2013	July.26	78.8	22.7	7.5	95.5	75.7	19.6
	2014	July.27	80.2	23.5	6.8	104.0	75.1	19.4
	^z Ave	July.26	78.0	22.4	7.8	97.8	76.1	17.1
			±2.69	±1.28	±1.12	±5.46	±1.31	±0.25
Superjami	2012	Aug.29	76.8	19.3	10.8	104.2	83.4	25.9
	2013	Aug.28	74.8	19.7	10.5	115.1	83.0	26.3
	2014	Aug.26	73.0	20.6	10.4	111.3	82.1	26.4
	^z Ave	Aug.28	74.9	19.9	10.6	110.2	82.8	26.2
			±1.9	±0.7	±0.2	±5.5	±0.7	±0.3
Superjami 2	2012	Aug.29	93.7	20.1	10.6	137.5	82.3	27.5
	2013	Aug.29	112.5	20.5	8.0	139.4	83.4	32.5
	2014	Sep.3	111.8	20.7	7.0	144.4	80.2	30.4
	^z Ave	Aug.30	106.0	20.4	8.5	140.4	82.0	30.1
			±10.7	±0.3	±1.9	±3.6	±1.6	±2.5

^z Values are expressed as means±SD.

2) 도정 및 품질특성

정현비율은 81.9%로 ‘흑진주벼’와 비슷하고, ‘슈퍼자미2호’의 현미 장폭비는 2.08의 장원형으로 ‘흑진주벼’에 비해 길이와 폭이 각각 12%, 26%로 증가되었다(표 11). 현미의 미량원소 함량에서 ‘슈퍼자미2호’가 ‘흑진주’에 비해 K, Ca 함량은 낮게 나타났고, 성분함량 분석 결과 단백질, 회분 함량은 낮게 열량, 지방, 탄수화물 함량은 비슷한 수준으로 나타났다(표 12).

Table 11. Grain appearance and milling properties of Superjami 2

Cultivar	Rough rice(mm)				Brown rice(mm)				Dehulling recovery(Brown/Rough) (%)
	Length	Width	Thickness	L/W ratio	Length	Width	Thickness	L/W ratio	
Heugjinju (a)	7.63	3.11	2.02	2.45	5.77	2.46	1.79	2.36	81.0
Superjami (b)	8.44	3.90	2.16	2.16	5.90	3.15	2.03	1.87	83.0
Superjami 2(c)	8.61	4.02	2.17	2.14	6.48	3.11	1.91	2.08	82.0
(c) / (a)	1.13	1.29	1.07		1.23	1.26	1.07		

z Values are expressed as means±SD.

Table 12. Varietal variation of general components.

Cultivar	Ca (mg/100g)	K (mg/100g)	Calorie (Kcal/100g)	Fat (g/100g)	Protein (g/100g)	Ash (g/100g)	Carbohydrate (g/100g)
Heugjinju	18.28±0.74	331.48±8.3	365	2.35±0.9	8.29±1.3	1.7±0.01	77.6
Superjami	13.55±0.35	271.40±4.9	351	2.06±0.6	9.27±1.8	1.4±0.04	75.4
Superjami2	11.29±0.27	305.39±4.2	361	2.04±0.5	9.46±1.2	1.5±0.04	77.8

z Values are expressed as means±SD.

또한 '슈퍼자미2호' 품종은 C3G 천연색소 함량이 현미 100g당 2013년에는 1,782 mg, 2014년에는 1,980 mg으로 흑진주보다 9배 이상 높은 것으로 확인되었다(그림 28).

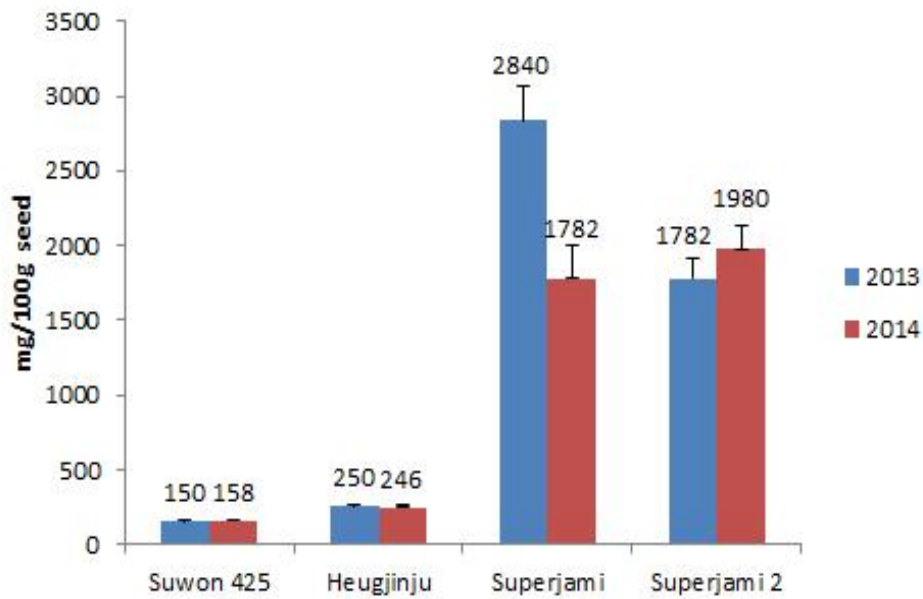


Fig. 28. Concentration of cyanidin 3-glucoside content in 'Superjami 2'

V. 결론

유전분석을 위해 사용된 C3GHi/흑진주 그리고 C3GHi/수원 425 교배 집단은 흑자색/흑자색 종피 조합으로 후대에서 다양한 색깔 변이를 보였다. 일부의 후대는 C3G 함량이 1,800mg/100g seed을 보이기도 하였다. 전체적으로 0 ~ 500 mg/100g seed 의 후대가 많았으며 히스토그램으로 보았을 때 정규분포에서 왼쪽으로 치우친 경향을 보였다. 여교배 특성상 한쪽 부모의 유전자형에 의해 치우침이 생길 수 있고, 교배조합의 모본 부분인 흑진주와 수원425가 각 각 인디카형, 자포니카 형으로 다양한 대립유전자 구성으로 인해 초월분리 현상이 일어난 것으로 생각되어진다. 기초 농업형질인 간장, 수장 그리고 수수의 분포도는 정규분포에 가까운 연속변이를 보였다. GBS를 이용한 SNP 마커 동정으로 총 6,703개의 마커가 발굴되었다. 이중 유효하고 의미 있는 마커 475개를 동정하여 물리적지도와 연관지도를 작성하였다. 연관지도 작성으로 17개의 연관군을 보였고 473개 마커가 위치했으며 마커 간 평균 거리는 3.55cM 이었다. QTL 분석 프로그램을 이용한 결과 간장과 관련된 QTL은 6번 염색체에 1개, 7번 염색체에 2개, 9번 염색체에 1개씩 총 4개 QTL이 탐색되었으며, 수장과 관련된 QTL은 1번 염색체와 11번 염색체에 총 2개의 QTL이 탐색되었고 수수와 관련된 QTL은 3번 염색체, 6번 염색체, 9번 염색체에 1개씩 그리고 10번 염색체에 2개로 총 5개 위치가 탐색되었다. C3G 함량 관련 QTL은

5번 염색체와 9번 염색체에 위치하였는데 각각 23.4%와 15.8%의 결정력을 보였다. 5번 위치 QTL 에는 종자 두께와 관련된 QTL이 보고된 바 있으며 세포벽 관련 유전자인 *EXPA3* 유전자와 세포크기를 작게하여 세포수를 증가시키는 *smos1* 유전자가 보고된 바 있다. Single-locus ANOVA를 통한 C3G함량과 마커간의 연관성 분석 결과로는 염색체 2번, 5번, 8번, 9번 10번 그리고 12번에 P값이 0.05 보다 작아 유의한 부분이 탐색되었다. 이 중 2번과 9번 위치는 유전자좌의 기여도가 상당히 높게 분석되었다. C3G 함량 관련 6개 QTL 과 실제 C3G 함량과의 관계를 살펴보았을 때 6개 모두를 가지고 있는 계통의 C3G 함량이 가장 높았으며 대체적으로 QTL 보유 개수가 많은 계통이 높은 C3G 함량을 보였다. 그러나 QTL 개수가 같은 계통내에서도 C3G 함량의 변이가 나타나는 것으로 보았을 minor QTL 들에 의한 함량 변이가 발생하는 것으로 추측된다. 상용품종 150 품종에 대해 9번 염색체 부위인 QTL내의 유전자 haplotype 분석을 한 결과 생태형을 기준으로 한 통일형과 자포니카형의 생태형으로 구분되었으며 일부는 출수기와 관련성이 높아 보였다. C3G 함량에 기여도가 높은 2번과 9번 염색체 QTL 내에서는 색소 생합성 전구물질합성에 관련되거나 색소의 안정성에 관여하는 유전자들이 발견되었다.

VI. 종합고찰

벼 종자의 종피색과 관련된 유전자는 *Pb*(chr. 4)와 *Pp*(chr. 1)가 흑자색 종피, *Rd*(chr. 7)와 *Rd*(chr. 1)가 적갈색 종피로 밝혀져 있다. 그러나, 유색미와 유색미 교배조합 혹은 유색미와 백미 교배조합 후대에서 다양한 변이를 보이고 안토시아닌 함량은 교배모본보다 월등히 높은 계통들이 선발 되었다.

Koes et al.(2005) 은 세포조직과 환경에 대한 반응으로 작용하는 MBW(MYB-bHLH-WDR) 단백질 복합체가 flavonoid 생합성 과정을 조절한다고 보고하여 다양한 구조유전자와 조절유전자가 색소 생합성에 관여한다고 보고하였다. 다른 연구에서는 135K *Oryza sativa* microarray를 이용하여 종자에서 특이적으로 발현하는 transcription factor을 연구하여 137개의 transcription factor 중 37%가 MYB 와 GT transcription factor group으로 분류되는 것을 보고하였다(Kim et al. 2011).

고전육종학에서 벼 종피의 흑자색은 *Pb*와 *Pp* loci에 의해 결정된다고 연구되었다(Nagao and Takahashi 1947). 그러나 이는 단순히 벼 종피에 색소의 착색여부만 설명이 되었고, 교배 조합후대의 초월분리 현상을 설명하기는 부족하였다. 벼 종피 색소의 함량에 영향을 주는 요인은 벼 종피 착색에 관여하는 *Pb*, *Pp* 에 위치한 대립유전자들의 상호작용 외에도 많은 종류의 transcription factor 들이 환경과의 상호작용 또는 세포조직 내의 발현 양상과 관련이 있을 것으로 생각되어진다.

2번 염색체 안에 위치한 *anthocyanin5-aromatic acyltransferase* (*Os02t0289500*) 유전자는 과남풀(*Gentiana triflora*) 꽃잎 색소를 중성에서 약염기성 사이 환경에서에서 안정화를 증가시키는 작용을 한다고 보고되었다(Fujiwara *et al.* 1997). 9번 염색체 QTL 안에 위치한 *Prephenate dehydratase* (*Os09t0566050*) 유전자는 식물에서 phenylalanine 아미노산의 생합성 마지막 과정에 관여하는 유전자로 알려져 있으며(Cotton RG. and Gibson F. 1965), phenylalanine은 식물의 2차대사산물인 flavonoid 생합성의 시작 아미노산이다.

Kim 등(2007)은 색소생합성 초기에 관여 하는 PAL 유전자는 상대적으로 후에 관여하는 DFR과 ANS 가 잎이나 종자에서 대량 발현되는데 비해 줄기나 뿌리에서 더 높게 발현되는 것을 보고 하여 PAL 유전자는 안토시아닌 생합성에만 관여하기 보다는 lignin의 생합성에도 관여하는 것으로 보고하였다. Raes 등(2003)도 조직 특이적으로, 시기적으로 생합성 유전자의 발현양이 조절되는 것을 보고하며 조절유전자의 역할이 중요하다고 하였다. 특히 세포질에서 일어나는 반응 초기 생합성 유전자들 보다는 후기 생합성 유전자 혹은 색소의 세포내 이동에 관여하는 유전자등에 의해 많이 관여된다고 보고하였다. 특히 uridine diphosphate-glucose:flavonoid 3-O-glucosyltransferase(UFGT)는 중요한 조절유전자로 보고되었는데(Kobayashi *et al.* 2002, 2004), 벼에서는 아직 발견되지 않았다(Bowles *et al.* 2005). 근래의 돌연변이개체를 이용한 연구 등에서는(Lepiniec L.

et al. 2006) 일명 *transparent testa(tt)*를 이용하여 MBW(MYB-bHLH-WDR) 복합체와 post-translational 조절 작용이 밝혀지고 있다(Wenjia X. *et al.* 2015).

종피의 색소 함량은 색소의 생합성에 관여하는 유전자뿐 아니라 조직 특이적으로 관여하는 조절유전자와 환경에 반응하는 신호유전자 그리고 물질의 이동이나 안정성 등에 관여하는 여러 유전자들이 복합적으로 작용하는 것으로 추론된다. Kim 등(2007)에 의하면 색소합성 관련 후기유전자들의 등숙기 온도에 따른 발현양을 비교하였을 때 21°C에서 보다 27°C에서 발현양이 2배 많아졌으며 탐색된 C3G 함량관련 QTL 부위의 Haplotype 분석 결과 자포니카와 통일형의 생태형 구분이 되었고, 일부는 출수기와 관련성이 높게 나타났다.

이러한 결과들을 종합적으로 살펴보면 벼의 종피 내 색소함량은 색소 생합성 관련 유전자들의 영향은 물론 이러한 생합성 관련 유전자들의 조직 특이적으로 발현이나 외부 환경과의 반응, 세포내 이동 및 저장 등과 관련된 조절유전자와 조절기작 등이 복잡하게 관여 되어 있는 것으로 여겨진다. 본 연구가 이러한 복잡한 현상을 해석하는데 디딤돌이 되기를 기대하며 이를 바탕으로 심도 있는 연구가 지속되기를 바란다.

VII. 참고문헌

Adaum-Blondon AF., Seignac M., Bannerot H. and Dron M. 1994. SCAR, RAPD and RFLP markers linked to a dominant gene(Are) conferring resistance to anthracnose in common bean. Theor. Appl. Genet. 88(6-7):865-870

Bardakci F. 2001. Random amplified polymorphic DNA(RAPD) markers. Turk. J. Biol. 25:185-196

Begum H., Spindel JE. Lalusin A., Borromeo T., Gregorio G. *et al.* 2015. Genome-wide association mapping for yield and other agronomic traits in an elite breeding population of tropical rice(*Oryza sativa*). PLoS One 10:e0119873

Borevitz JO., Xia Y., Blount J., Dixon RA. and Lamb C. 2000. Activation tagging identifies a conserved MYB regulator of phenylpropanoid biosynthesis. Plant Cell 12: 2383-2393.

Botstein D., White RL., Skolnick M. and Davis RW. 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am J. Hum. Genet. 32(3):314-333

Bovy A., de Vos R., Kemper M., Schijlen E., Almenar Pertejo M., Muir S., Collins G., Robinson S., Verhoeyen M., Hughes S., Santos-Buelga C. and van Tuinen A. 2002. High-flavonol tomatoes resulting from the heterologous expression of the maize transcription factor genes LC and C1. Plant Cell. 14:2509-2526.

Bowles D., Lsayenkova J., Lim EK. and Poppenberger B. 2005. Glucosyl-transferases: Managers of small molecules. Curr. Opin.

Plant Biol. 8:254-263

Burrell AM., Repper AE., Hodnett G., Goolsby JA., Overholt WA. *et al.* 2015. Exploring origins, invasion history and genetic diversity of *Imperata cylindrica*(L.) P. Beauv.(Cogongrass) in the United States using genotyping by sequencing. *Mol. Ecol.* 24(9):2177-2193

Chelkowski J. and Stêpieñ L. 2001. Molecular markers for leaf rust resistance genes in wheat. *J. Appl. Genet.* 42(2):117-126

Ching ADA., Caldwell KS., Jung M., Dolan M., Smith OS., Tingey S., Morgante M. and Rafalski A. 2002. SNP frequency, haplotype structure and linkage disequilibrium in elite maize inbred lines. *BMC Genet.* 3:19

Choi Y., Kim MH., Shin JJ., Park JM. and Lee J. 2003. The antioxidant activities of the some commercial teas. *J. Korean Soc. Food Sci. Nutr.* 32:723-727

Cotton RG. and Gibson F. 1965. The Biosynthesis of Phenylalanine and Tyrosine; Enzymes Converting Chorismic acid into Prephenic acid and their Relationships to Prephenate Dehydratase and Prephenate Dehydrogenase. *Biochim. Biophys. Acta.* 100:76-88

Davey JW., Hohenlohe PA., Etter PD., Boone JQ., Catchen JM. *et al.*, 2011. Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nat. Rev. Genet.* 12(7):449-510

Dai J., Gupet A., Gates L. and Mumper RJ. 2009. A

comprehensive study of anthocyanin-containing extracts from selected blackberry cultivars: Extraction methods, stability, anticancer properties and mechanisms. *Food and Chemical Toxicology* 47:837~847

Di Carlo M., Mascolo GN., Izzo AA. and Capasso F. 1999. Flavonoids: old and new aspects of a class of natural therapeutic drugs. *Life Sci.* 65(4):337-353

Elmer I. Humira S. and Francois B. 2015. Association mapping of QTLs for sclerotinia stem rot resistance in a collection of soybean plant introductions using a genotyping by sequencing (GBS) approach. *BMC Plant Biol.* 15:5

Fang YZ., Yang S. and Wu G. 2002. Free radicals, antioxidants and nutrition. *Nutrition* 18(10):872-879

Fujiwara H., Tanaka Y., Fukui Y., Nakao M., Ashikari T., Kusumi T. 1997. Anthocyanin 5-aromatic acyltransferase from *Gentiana trilora* purification, characterization and its role in anthocyanin biosynthesis. *Eur. J. Biochem.* 249:45-51

Furuta T., Ashikari M., Jena KK., Doi K. and Reuscher S. 2016. Adapting Genotyping-by-Sequencing for Rice F2 Populations. *G3* 7(3):881-893

Gordon MH. 1996. Dietary antioxidants in disease prevention. *Nat. Pro. Rep.* 13:265-273

Guo W., Zhang T., Shen X., Yu JZ. and Kohel RJ. 2003. Development of SCAR marker linked to a major QTL for high

fiber strength and its usage in molecular-marker assisted selection in upland cotton. *Crop Sci.* 43:2252-2256

Ham TH., Kwon SW. and Ryu SN. 2015. A Rice Variety 'Superjami 2', with Large-Grain and High Contents of Cyanidin 3-glucoside. *Korean J. Breed. Sci.* 47(3):299-305

Han SJ., Kim JS., Chae SW., Ryu SN., Hyun JW., Son KH., Sohn HY., Chang HW. and Kang SS. 2004. Biological Screening of Extracts from the Colored Rice Cultivars. *Kor. J. Pharmacogn.* 35(4):346~349

Han SJ., Trinh HT., Hong SS., Ryu SN. and Kim DH. 2007. Antipruritic Effect of Black Colored Rice. *Natural Product Science* 13(4):373~377

Han SJ., Ryu SN., Trinh HT., Joh EH., Jang SY., Han MJ. and Kim DH. 2009. Metabolism of cyanidin-3-*O*-beta-D-glycoside isolated from black rice and its anti-scratching behavioral effect in mice. *J. of Food Science* 74:H253-H258

Heller W. and Forkmann G. 1993. Biosynthesis of flavonoids. In:Harborne JB. (Ed.), *The Flavonoids: Advances in Research since 1986*. Chapman & Hall, London, pp. 499-535.

Hemmat M., Weeden NF., Manganaris AG. and Lawson DM. 1994. Molecular marker linkage map for apple. *J. Heredity* 85:4-11

Hollman PCH., Hertog MGL. and Katan MB. 1996. Role of dietary flavonoids in protection against cancer and coronary heart disease. *Bioact. Compd. Food* 24(3):785-789

Huang J. and Sun M. 1999. A modified AFLP with fluorescence-labelled primers and automated DNA sequencer detection for efficient fingerprinting analysis in plants. *Biotechnol. Techn.* 14:277-278

Huang YF., Poland JA., Wight CP., Jackson EW. and Tinker NA., 2014. Using genotyping-by-sequencing (GBS) for genomic discovery in cultivated oat. *PLoS One* 9:e102448

Hsieh SC. and Chang TM. 1964. Genic analysis in rice IV. Genes for purple pericarp and other characters. *Jap. J. Breeding* 14(3):1-9

Jayaprakasha GK., Singh RP. and Sakariah KK. 2001. Antioxidant activity of grape seed(*Vitis vinifera*) extracts on peroxidation models in vitro. *Food Chem.* 73(3):285-290

Kenkt P., Kumpulainen J., Javinen R., Rissanen H., Heliovaara M., Reunanen A., Hakulinen T. and Aromaa A. 2002. Flavonoid intake and risk of chronic diseases. *Am. J. Clin. Nutr.* 76(3):560-568

Kim CK., Cho MA., Choi YH., Kim JA., Kim YH., Kim YK. and Park SH. 2011. *J. Appl. Genetics* 52:161-169

Kim HY., Kim JH., Lee SA., Ryu SN., Han SJ. and Hong SG. 2010. Antioxidative and anti-diabetic activity of C3GHi, novel black rice breed. *Kor. J. Crop Sci.* 55(1):38~46

Kinoshita T. 1984. Gene analysis and linkage map. *Biology of Rice*:187-274

Kinoshita T. 1995. Report of Committee on Gene Symbols, Nomenclature and Linkage group. RGN 12:9-153

Kiss GB., Csanádi G., Kálman K., Kaló P. and Ôkresz L. 1993. Construction of a basic linkage map for alfalfa using RFLP, RAPD, isozyme and morphological markers. Mol. Gen. Genet. 238:129-137

Kobayashi S., Ishimaru M., Hiraoka K. and Honda C. 2002. *Myb*-related genes of the Kyoho grape(*Vitis labruscana*) regulate anthocyanin biosynthesis. Planta 215:924-933

Kobayashi S., Gotto-Yomamoto N. and Hirochika H. 2004. Restrotransposon-induced mutations in grape skin color. Science 304:982

Konieczny A. and Ausubel FM. 1993. A procedure for mapping Arabidopsis mutations using co-dominant ecotype-specific PCR-based markers. Plant J. 4:403-410

Kuhn MA. 2003. Oxygen free radicals and antioxidants: An overview of how antioxidants protect the body from disease. Am. J. Nurs. 103(4):58-62

Lepiniec L., Debeaujon I., Routaboul JM., Baudry A., Pourcel L., Nesi N. and Caboche M. 2006. Genetics and biochemistry of seed flavonoids. Annu. Rew. Plant Biol. 57:405-430

Lin M., Cai S., Wang S., Liu S., Zang G. *et al.* 2015. Genotyping-by-sequencing(GBS) identified SNP tightly linked to QTL for pre-harvest sprouting resistance. Theor. Appl. Genet.

128:1385-1395

Litt M. and Luty JA. 1986. A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. *Am. J. Hum. Genet.* 44:397-401

Lloyd AM., Walbot V. and Davis RW. 1992. Arabidopsis and Nicotiana anthocyanin production activated by maize regulators R and C1. *Science* 258:1773-1775

Martin GB., Williams JGK. and Tanksley SD. 1991. Rapid identification of markers linked to a *Pseudomonas* resistance gene in tomato by using random primers and near-isogenic lines. *Proc. Natl. Acad. Sci.* 88:2336-2340

Matsui K., Tanaka H. and Ohme-Takagi M. 2004. Suppression of the biosynthesis of proanthocyanidin in Arabidopsis by a chimeric PAP1 repressor. *Plant Biotechnol. J.* 2:1-7

McDermott JM., Brandle U., Dutly F., Haemmerli UA., Keller S., Muller KE. and Wolf MS. 1994. Genetic variation in powdery mildew of barley: development of RAPD, SCAR and VNTR markers. *Phytopathology* 84:1316-1321

Mullis KB. and Faloona F. 1987. Specific synthesis of DNA in vitro via polymerase chain reaction. *Methods Enzymol.* 155:350-355

Osawa T., Ramarathnam N., Kawakishi S. and Namiki M. 1992. Phenolic Compounds in Food and Their Effects on Health II. Antioxidants & Cancer Prevention. Washington, American

Chemical Society

Oh SK., Choi HC., Cho MY. and Kim SU. 1996. Extraction method of anthocyanin and tannin pigments in colored rice. *Agri. Chem. and Biotech.* 39(4):327-331

Owens DK., Alerding AB., Crosby KC., Bandara AB., Westwood JH., Winkel BS. 2008. Functional analysis of a predicted flavonol synthase gene family in *Arabidopsis*. *Plant Physiol* 147:1046-1061

Paran I. and Michelmore RW. 1993. Development of reliable PCR-based markers linked to downy mildew resistance genes in lettuce. *Theor. Appl. Genet.* 85:985-993

Paran I., Kesseli R. and Michelmore R. 1991. Identification of restriction-fragment-length-polymorphism and random amplified polymorphic DNA markers linked to downy mildew resistance genes in lettuce, using near isogenic lines. *Genome* 34:1021-1207

Park YS., Kim SJ. and Chang HI. 2008, Isolation of anthocyanin from black rice (*Heugjinjubyeo*) and screening of its antioxidant activities. *Kor. J. Microbiol. Biotechnol.* 36(1):55-60

Quattrocchio F., Wing JF., Leppen H., Mol J. and Koes R. 1993. Regulatory genes controlling anthocyanin pigmentation are functionally conserved among plant species and have distinct sets of target genes. *The Plant Cell* 5:1497-1512

Quattrocchio F., Wing JF., van der Woude K., Mol J. and Koes R. 1998. Analysis of bHLH and MYB domain proteins: species-specific regulatory differences are caused by divergent

evolution of target anthocyanin genes. *The Plant Journal* 13:475-488

Rabbi IY., Hamblin MT., Kumar PL., Gedil MA., Ikpan AS. *et al.* 2014. High-resolution mapping of resistance to cassava mosaic geminiviruses in cassava using genotyping-by-sequencing and its implications for breeding. *Virus Res.* 186:87-96

Raes J., Rohde A., Christensen JH., van de Peer Y. and Boerjan W. 2003. Genome-wide characterization of the lignification toolbox in *Arabidopsis*. *Plant Physiol.* 133:1051-1071

Rafalski A. 2002. Applications of single nucleotide polymorphisms in crop genetics. *Curr. Opin. Plant Biol.* 5:94-100

Reddy VS., Dash S. and Reddy AR. 1995. Anthocyanin pathway in rice (*Oryza sativa* L.): Identification of a mutant showing dominant inhibition of anthocyanins in leaf and accumulation of proanthocyanidins in pericarp. *Ther. Appl. Genet.* 91:301-312

Ross JA. and Kasum CM. 2002. Dietary Xavonoids: bioavailability, metabolic effects, and safety. *Annu. Rev. Nutr.* 22:19-4

Ryu SN., Park SZ. and Ho CT. 1998. High performance liquid chromatographic determination of anthocyanin pigments in some varieties of black rice. *Journal of Food and Drug analysis* 6(4):729-736

Ryu SN. 2014. *Science of Functional Rice*, KNOU Press

Sachidanandam R., Weissman D., Schmidt SC., Kakol JM., Stein

LD., Marth G., Sherry S., Mullikin JC., Mortimore BJ. and Willey DL. 2001. A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature* 409:928-933

Saghai Maroof MA., Biyashev RM., Yang GP., Zhang Q. and Allard RW. 1994. Extraordinarily polymorphic microsatellite DNA in barley: species diversity, chromosomal locations and population dynamics. *Proc. Natl. Acad. Sci.* 91:5466-5470

Schijlen EG., de Vos RCH., van Tunen AJ. and Bovy AG. 2004. Modification of flavonoid biosynthesis in crop plants. *Phytochemistry* 65:2631-2648.

Schlötterer C. and Tautz D. 1992. Slippage synthesis of simple sequence DNA. *Nucleic Acids Res.* 20:2211-2215

Semagn K., Bjornstad A. and Ndjioudjop MN. 2006. An overview of molecular marker methods for plants. *Afr. J. Biotech.* 5:2540-2568

Shin YM., Park HJ., Yim SD., Baek NI., Lee CH., Ahn G. and Woo YM. 2006. Transgenic rice lines expressing maize C1 and R-Sregulatory genes produce various flavonoids in the endosperm. *Plant Biotech. J.* 4:303-315.

Singh H., Deshmukh RK., Singh A., Singh AK. and Gaikwad K. 2010. Highly variable SSR markers suitable for rice genotyping using agarose gels. *Mol. Breed.* 25:359-364

Spindel J., Wright M., Chen C., Cobb J., Gage J. *et al.* 2013. Bridging the genotyping gap: using genotyping by sequencing

(GBS) to add high-density SNP markers and new value to traditional bi-parental mapping and breeding populations. Theor. Appl. Genet. 126:2699-2716

Tahara S. 2007. A journey of twenty-five years through the ecological biochemistry of flavonoids. Biosci. Biotechnol. Biochem. 71:1387-404.

Tautz D. and Renz M. 1984. Simple sequences are ubiquitous repetitive components of eukaryotic genomes. Nucleic Acids Res. 12:4127-4138

Tohge T., Nishiyama Y., Hirai MY. *et al.* 2005. Functional genomics by integrated analysis of metabolome and transcriptome of Arabidopsis plants over-expressing an MYB transcription factor. Plant J. 42:218-235.

Torress AM., Weeden NF. and Martin A. 1993. Linkage among isozyme, RFLP and RAPD markers. Plant Physiol. 101:394-452

van de Wiel C., Arens P. and Vosman B. 1999, Microsatellite retrieval in lettuce(*Lactuca sativa* L.). Genome 42:139-149

Verhoeven ME., Bovy A., Collins G., Muir S., Robinson S., de Vos CHR. and Colliver S. 2002. Increasing antioxydant levels in tomatoes through modification of the flavonoid biosynthetic pathway. J. Exp. Bot. 53:2099-2106.

Verlangieri A., Kapeghian J., el-Dean JC. and Bush SM. 1985. Fruit and vegetable consumption and cardiovascular mortality. Med. Hypotheses. 16:7~15

Vos P., Hogers R., Bleeker M., Reijans M., van de Lee T., Hornes M., Frijters A., Pot J., Peleman J., Kuiper M. and Zabeau M. 1995. AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res.* 23:4407-4414

Wenz HM., Robertson JM., Menchen S., Oaks F., Demorest DM., Scheibler D., Rosenblum BB., Wike C., Gilbert DA. and Efcavitch JW. 1998. High-precision genotyping by denaturing capillary electrophoresis. *Genome Res.* 3:69-80

Williams JGK., Kubelik AR., Livak KJ., Rafalski JA. and Tingey SV. 1991. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Res.* 18:6531-6535

Williams RJ., Spencer JP. and Rice-Evans C. 2004. Flavonoids: antioxidants or signalling molecules? *Free Radic. Biol. Med.* 36:38-49.

Winkel-Shirley B. 2001. Flavonoid biosynthesis: a colorful model for genetics, biochemistry, cell biology and biotechnology. *Plant Physiol.* 126:485-493.

Wu KS. and Tanksley SD. 1993 Abundance, polymorphism and genetic mapping of microsatellites in rice. *Mol. Gen. Genet.* 241:225-235

Xu W. Dubos C. and Lepiniec L. 2015. Transcriptional control of flavonoid biosynthesis by MYB-bHLH-WDR complexes. *Trends in Plant Sci.* 20(3):176-185

Zhu YL., Song QJ., Hyten DL., van Tassel CP., Matukumalli LK.,

Grimm DR. *et al.* 2003, Single-nucleotide polymorphisms in soybean. *Genetics* 163:1123-1134

Identification of QTLs for High-Dense Cyanidin 3-Glucoside Content in Rice Grain

Ham, Tae-ho
Department of Plant Science
The Graduate School
Seoul National University

Among the constituents of anthocyanin, which is a main pigment of black seed coat color, C3G (Cyanidin 3-Glucoside), which has excellent antioxidative ability and has a high content of anthocyanin pigment. To find out the genes related C3G content, backcrossed combination was constructed that C3GHi(Heugjinju/S. 425) crossed with Heugjinju variety and Suwon 425 variety. The combination progressed to F₅ generation by SSD (Single Seed Descent) method from F₂ generation. C3G content was measured with culmn length, panicle length and number of panicle, which are basic agricultural traits. The C3GHi / Suwon 425 backcrossed RIL population was defined as the analysis group by comparing the population numbers and the

distribution of C3G content. The genetic analysis method was GBS (Genotyping by Sequencing) using NGS (Next Generation Sequencing) technology. A total of 6,703 SNPs were detected by the GBS method. A total of 475 significant SNP markers were screened by filtering and binning to eliminate uninformative or not properly segregated SNP markers in population. Based on the nucleotide sequence of 475 selected SNP markers, a physical map was created and a linkage map was constructed based on the SNP marker genotype. A total of 17 linkage groups were constructed. Chromosome 4, which had a large gap on the gingival margin, was divided into three groups, and chromosomes 1, 5, and 11 were divided into two groups. There were 473 markers in the association map, and the mean distance between the markers was 3.55 cM. Single-locus ANOVA was performed to association analysis between the SNP markers and the C3G content. Six QTLs associated with the C3G content on chromosome 2, 5, 8, 9, 10 and 12. The candidate gene was identified in chromosome 2 QTL by narrowing down the QTLs on chr. 2 and chr. 9 which had significant gene effects to C3G content in rice. Superjami 2 was developed by cross with C3GHi

and Daeripbyeo 1, blackish purple and white pericarp respectively. It has late maturing and tall height as compared with Heugjinju. In fuctional components, the microelement like calcium and potassium was little bit low to Heugjinju, but the content of C3G was higher than Heugjinju as 1,782 mg(100g seed weight) and 1,980 mg(100g seed weight) in 2013, 2014 year respectively. It was nine-times more than Heugjinju.

Keywords : Colored rice, Anthocyanin, Genotyping by Sequencing(GBS), Linkage Mapping, Quantitative Trait Loci(QTL)

Student Number : 2003-30358

VIII. 부록

chr.	bp	gene name	gene function
02	10,552,438	Os02g0282900	Similar to 68 kDa protein HP68.(Os02t0282900-01);Transcript variants=Os02t0282900-01;CGSNL Gene Sym
02	10,634,660	Os02g0283800	Similar to SERK1(Fragment). (Os02t0283800-01);Transcript variants=Os02t0283800-01;OryzabaseGene Sy
02	10,677,732	Os02g0284300	Similar to predictedprotein. (Os02t0284300-00);Transcript variants=Os02t0284300-00
02	10,693,427	Os02g0284500	FAR1 DNA binding domaindomain containing protein. (Os02t0284500-01);Transcriptvariants=Os02t028450
02	10,700,829	Os02g0284600	Similar to 60Sribosomal protein L27. (Os02t0284600-01);Transcriptvariants=Os02t0284600-01
02	10,715,073	Os02g0285300	DREPP plasma membranepolypeptide family protein. (Os02t0285300-01);Transcriptvariants=Os02t0285300
02	10,725,543	Os02g0285700	Similar to predictedprotein. (Os02t0285700-01);Transcript variants=Os02t0285700-01
02	10,739,333	Os02g0285775	Non-protein codingtranscript. (Os02t0285775-00);Transcript variants=Os02t0285775-00
02	10,739,451	Os02g0285800	Similar to GTP-bindingprotein typA (Tyrosine phosphorylated protein A).(Os02t0285800-01)%3BSimilar
02	10,747,495	Os02g0285900	No apical meristem(NAM) protein domain containing protein. (Os02t0285900-00);Transcriptvariants=Os
02	10,758,967	Os02g0286200	Lipase class 3family protein. (Os02t0286200-01);Transcript variants=Os02t0286200-01
02	10,764,855	Os02g0286350	Similar to Non-specificlipid-transfer protein. (Os02t0286350-00);Transcriptvariants=Os02t0286350-0
02	10,766,482	Os02g0286500	Similar to AT hookmotif family protein expressed. (Os02t0286500-00);Transcriptvariants=Os02t028
02	10,772,392	Os02g0286700	Similar toChlorophyllase (Fragment). (Os02t0286700-00);Transcriptvariants=Os02t0286700-00
02	10,778,463	Os02g0286850	Similar to Stripe rustresistance protein Yr10. (Os02t0286850-00);Transcriptvariants=Os02t0286850-0
02	10,787,293	Os02g0286891	Similar toOSIGBa0161P06.1 protein. (Os02t0286891-01);Transcriptvariants=Os02t0286891-01
02	10,822,643	Os02g0287000	Similar to 40Sribosomal protein S3a (CYC07 protein). (Os02t0287000-01) Ribosomal proteinsS3Ae dom
02	10,854,309	Os02g0287800	F-box domain Skp2-like domain containing protein. (Os02t0287800-00);Transcriptvariants=Os02t0287
02	10,857,586	Os02g0287900	Cyclin-like F-boxdomain containing protein. (Os02t0287900-01);Transcriptvariants=Os02t0287900-01;O
02	10,858,501	Os02g0287950	Non-protein codingtranscript. (Os02t0287950-00);Transcript variants=Os02t0287950-00
02	10,861,422	Os02g0288000	Cyclin-like F-boxdomain containing protein. (Os02t0288000-01);Transcriptvariants=Os02t0288000-01;O
02	10,864,069	Os02g0288100	Similar toPectinesterase (EC 3.1.1.11) (Fragment). (Os02t0288100-01);Transcriptvariants=Os02t02881
02	10,870,290	Os02g0288200	Protein of unknownfunction DUF573 domain containing protein. (Os02t0288200-01) Hypotheticalconse
02	10,880,647	Os02g0288400	Similar to MRP-like ABCtransporter. (Os02t0288400-00);Transcript variants=Os02t0288400-00;CGSNLGen
02	10,897,250	Os02g0288600	Similar to BURPdomain-containing protein 4. (Os02t0288600-01);Transcriptvariants=Os02t0288600-01;O

02	10,908,431	Os02g0288700	Similar to MRP-like ABCtransporter. (Os02t0288700-00);Transcript variants=Os02t0288700-00;CGSNLGen
02	10,908,608	Os02g0288722	Non-protein codingtranscript. (Os02t0288722-00);Transcript variants=Os02t0288722-00
02	10,931,592	Os02g0289000	BURP domain containingprotein. (Os02t0289000-01);Transcript variants=Os02t0289000-01
02	10,945,077	Os02g0289500	Similar to anthocyanin5-aromatic acyltransferase. (Os02t0289500-01);Transcriptvariants=Os02t0289500
02	10,960,662	Os02g0290000	Pentatricopeptiderepeat domain containing protein. (Os02t0290000-01);Transcriptvariants=Os02t02900
02	10,972,770	Os02g0290300	WD40/YVTN repeat-likedomain containing protein. (Os02t0290300-01);Transcriptvariants=Os02t0290300-
02	10,990,473	Os02g0290400	Ribosomal biogenesis%2Cmethyltransferase EMG1/NEP1 domaincontaining protein. (Os02t0290400-01)
02	10,994,252	Os02g0290500	Similar to Ubiquitin.(Os02t0290500-01) Similar to phosphatidylinositol 3- and 4-kinase familypro
02	10,999,890	Os02g0290600	Prolyl oligopeptidasefamily protein Abiotic stress tolerance (Os02t0290600-01)%3BSimilar topred
02	11,009,350	Os02g0290900	Similar to GDSL-motiflipase/hydrolase family protein. (Os02t0290900-00);Transcriptvariants=Os02t02
02	11,015,828	Os02g0291000	Calcineurin B(Fragment). (Os02t0291000-00);Transcript variants=Os02t0291000-00;CGSNL GeneSymbol=CB
02	11,057,896	Os02g0291400	Similar to CalcineurinB-like protein 8. (Os02t0291400-00);Transcriptvariants=Os02t0291400-00;CGSNL
02	11,071,696	Os02g0291600	Lipase%2C GDSL domaincontaining protein. (Os02t0291600-01);Transcriptvariants=Os02t0291600-01;CGSN
02	11,116,177	Os02g0292600	Lipase%2C GDSL domaincontaining protein. (Os02t0292600-01);Transcriptvariants=Os02t0292600-01;CGSN
02	11,126,248	Os02g0292800	Protein of unknownfunction DUF599 family protein. (Os02t0292800-01);Transcriptvariants=Os02t029280
02	11,151,750	Os02g0293300	Similar to Gene X-likeprotein (Fragment). (Os02t0293300-01);Transcriptvariants=Os02t0293300-01
02	11,157,096	Os02g0293400	Similar to Ubiquitinligase SINAT5 (EC 6.3.2.-) (Seven in absentia homolog 5). Splice isoform 2.(Os
02	11,162,557	Os02g0293500	Similar to CLP proteaseregulatory subunit CLPX precursor. (Os02t0293500-01);Transcriptvariants=Os0
02	11,183,598	Os02g0293900	Similar to nucleic acidbinding protein. (Os02t0293900-00);Transcript variants=Os02t0293900-00
02	11,186,713	Os02g0294000	Micro-fibrillar-associated 1 C-terminaldomain containing protein. (Os02t0294000-00);Transcript v
02	11,188,432	Os02g0294100	Cyclin-like F-boxdomain containing protein. (Os02t0294100-01)Cyclin-like F-box domaincontaining
02	11,199,426	Os02g0294600	Similar to transducinfamily protein / WD-40 repeat family protein. (Os02t0294600-01);Transcriptvar
02	11,209,135	Os02g0294700	Topoisomerasell-associated protein PAT1 domain containing protein.(Os02t0294700-01);Transcript var
02	11,292,722	Os02g0295700	Aldehydedehydrogenase conserved site domain containing protein.(Os02t0295700-01)%3BAldehyde deh
02	11,339,961	Os02g0296350	Non-protein codingtranscript. (Os02t0296350-01);Transcript variants=Os02t0296350-01
02	11,340,114	Os02g0296400	Similar to RCD1.(Os02t0296400-00);Transcript variants=Os02t0296400-00;Oryzabase Gene SymbolSynonym
02	11,353,288	Os02g0296800	Cobalamin (vitamin B12)biosynthesis CbiX domain containing protein. (Os02t0296800-01);Transcriptva
02	11,362,531	Os02g0297132	Similar to heat-shockprotein 101. (Os02t0297132-00);Transcript variants=Os02t0297132-00

02	11,376,851	Os02g0297200	Protein of unknown function DUF594 family protein. (Os02t0297200-01);Transcriptvariants=Os02t029720
02	11,399,635	Os02g0297600	Protein of unknown function DUF707 family protein. (Os02t0297600-01)Protein of unknownfunction D
02	11,414,988	Os02g0297900	Similar to F-box domain containing protein expressed. (Os02t0297900-00);Transcriptvariants=Os02t
02	11,427,608	Os02g0298200	Similar to Resistance protein candidate (Fragment). (Os02t0298200-01);Transcriptvariants=Os02t02982

chr.	bp	gene name	gene function
05	24,913,012	Os05g0505383	Hypothetical gene. (Os05t0505383-01)
05	24,961,645	Os05g0505692	Conserved hypothetical protein. (Os05t0505692-01)
05	24,987,033	Os05g0505900	Similar to CLE family OsCLE507 protein. (Os05t0505900-01)
05	24,987,033	Os05g0505900	Similar to CLE family OsCLE507 protein. (Os05t0505900-01)
05	24,997,324	Os05g0506000	Similar to MS5-like protein (Fragment). (Os05t0506000-01)
05	25,007,029	Os05g0506200	KOW domain containing protein. (Os05t0506200-01)
05	25,013,673	Os05g0506300	Conserved hypothetical protein. (Os05t0506300-01)
05	25,013,673	Os05g0506300	Conserved hypothetical protein. (Os05t0506300-01)
05	25,022,159	Os05g0506450	Similar to esterase. (Os05t0506450-00)
05	25,022,958	Os05g0506500	Conserved hypothetical protein. (Os05t0506500-00)
05	25,025,141	Os05g0506600	Lipase C GDSL domain containing protein. (Os05t0506600-00)
05	25,035,919	Os05g0506800	Hypothetical conserved gene. (Os05t0506800-01)
05	25,041,531	Os05g0506900	Similar to Brix domain-containing protein 1. (Os05t0506900-01)
05	25,044,642	Os05g0507000	Similar to JHL23J11.5 protein. (Os05t0507000-01)
05	25,055,345	Os05g0507300	Similar to Calreticulin-3. (Os05t0507300-01)
05	25,063,389	Os05g0507400	Conserved hypothetical protein. (Os05t0507400-01)
05	25,153,902	Os05g0508300	Similar to Papain-like cysteine peptidase XBCP3. (Os05t0508300-01)
05	25,159,955	Os05g0508400	Mannose-binding lectin domain containing protein. (Os05t0508400-01)
05	25,164,335	Os05g0508500	Protein of unknown function DUF594 domain containing protein. (Os05t0508500-01)
05	25,164,335	Os05g0508500	Protein of unknown function DUF594 domain containing protein. (Os05t0508500-01)
05	25,176,651	Os05g0508700	TRAF-type domain containing protein. (Os05t0508700-01)
05	25,176,651	Os05g0508700	TRAF-type domain containing protein. (Os05t0508700-01) Similar to Math domain containing protein (Fragment). (Os05t0508700-02)
05	25,176,651	Os05g0508700	TRAF-type domain containing protein. (Os05t0508700-01) Similar to Math domain containing protein (Fragment). (Os05t0508700-02)
05	25,177,260	Os05g0508750	Hypothetical protein. (Os05t0508750-00)
05	25,188,258	Os05g0508800	Similar to H1flk (Fragment). (Os05t0508800-01)
05	25,193,479	Os05g0508900	Similar to cDNA clone:001-011-D2C full insert sequence. (Os05t0508900-01)
05	25,199,743	Os05g0509101	Conserved hypothetical protein. (Os05t0509101-00)

05	25,214,532	Os05g0509200	Thioredoxin fold domain containing protein. (Os05t0509200-01)
05	25,218,662	Os05g0509300	Pseudouridine synthase TruD domain containing protein. (Os05t0509300-01)
05	25,224,888	Os05g0509400	Similar to DNA binding protein-like. (Os05t0509400-01)
05	25,229,702	Os05g0509500	Similar to Signal recognition particle 54 kDa protein 2 (SRP54). (Os05t0509500-01)
05	25,254,990	Os05g0509700	Nucleic acid-binding OB-fold domain containing protein. (Os05t0509700-01)
05	25,255,960	Os05g0509800	Protein of unknown function DUF1677 C plant domain containing protein. (Os05t0509800-00)
05	25,260,687	Os05g0509900	Winged helix repressor DNA-binding domain containing protein. (Os05t0509900-01)
05	25,260,687	Os05g0509900	Winged helix repressor DNA-binding domain containing protein. (Os05t0509900-01)
05	25,260,687	Os05g0509900	Winged helix repressor DNA-binding domain containing protein. (Os05t0509900-01)
05	25,260,687	Os05g0509900	Winged helix repressor DNA-binding domain containing protein. (Os05t0509900-01)
05	25,260,687	Os05g0509900	Winged helix repressor DNA-binding domain containing protein. (Os05t0509900-01)
05	25,266,272	Os05g0510100	Protein of unknown function DUF567 family protein. (Os05t0510100-01)
05	25,267,915	Os05g0510200	Similar to Glycine-rich cell wall structural protein 1.0 precursor (GRP 1.0). (Os05t0510200-01)
05	25,274,418	Os05g0510300	Peptidase C19 ubiquitin carboxyl-terminal hydrolase 2 family protein. (Os05t0510300-01)
05	25,283,108	Os05g0510400	Cyclin-like F-box domain containing protein. (Os05t0510400-01)
05	25,283,108	Os05g0510400	Cyclin-like F-box domain containing protein. (Os05t0510400-01)
05	25,291,060	Os05g0510600	Similar to tyrosine/DOPA decarboxylase 1. (Os05t0510600-01)
05	25,295,858	Os05g0510700	Similar to BSD domain containing protein. (Os05t0510700-01)
05	25,300,845	Os05g0510800	Similar to Cellulose synthase-like family C1 protein. (Os05t0510800-01)
05	25,303,628	Os05g0510932	Conserved hypothetical protein. (Os05t0510932-00)
05	25,312,775	Os05g0511000	TRAM LAG1 and CLN8 homology domain containing protein. (Os05t0511000-01)
05	25,312,775	Os05g0511000	TRAM LAG1 and CLN8 homology domain containing protein. (Os05t0511000-01)
05	25,317,044	Os05g0511050	Similar to aspartyl protease family protein. (Os05t0511050-00)
05	25,321,363	Os05g0511300	N2227-like domain containing protein. (Os05t0511300-01)
05	25,331,884	Os05g0511400	Similar to Protein kinase GhCLK1 (Fragment). (Os05t0511400-01)

05	25,343,800	Os05g0511500	Similar to Lipoic acid synthase-like protein. (Os05t0511500-01)
05	25,346,821	Os05g0511700	Putative rRNA methylase family protein. (Os05t0511700-01)
05	25,350,274	Os05g0511800	Protease inhibitor I4 serpin domain containing protein. (Os05t0511800-00)
05	25,357,253	Os05g0512000	Zinc finger RING/FYVE/PHD-type domain containing protein. (Os05t0512000-01)
05	25,379,151	Os05g0512050	Conserved hypothetical protein. (Os05t0512050-00)
05	25,379,290	Os05g0512100	Uncharacterised conserved protein UCP031279 domain containing protein. (Os05t0512100-00)
05	25,382,364	Os05g0512200	Similar to T4P13.26 protein. (Os05t0512200-01)
05	25,387,145	Os05g0512301	Similar to CLE family OsCLE508 protein. (Os05t0512301-00)
05	25,390,732	Os05g0512400	IQ motif%2C EF-hand binding site domain containing protein. (Os05t0512400-01)
05	25,392,886	Os05g0512500	CT20 family protein. (Os05t0512500-01)
05	25,399,448	Os05g0512600	X8 domain containing protein. (Os05t0512600-01)
05	25,436,359	Os05g0513100	Similar to PICYC4 protein (Fragment). (Os05t0513100-01)
05	25,446,943	Os05g0513200	NLI interacting factor domain containing protein. (Os05t0513200-00)
05	25,453,751	Os05g0513300	Hypothetical conserved gene. (Os05t0513300-00)
05	25,456,387	Os05g0513350	Similar to predicted protein. (Os05t0513350-00)
05	25,458,603	Os05g0513400	Similar to non-imprinted in Prader-Willi/Angelman syndrome region protein 1. (Os05t0513400-01)
05	25,458,603	Os05g0513400	Similar to non-imprinted in Prader-Willi/Angelman syndrome region protein 1. (Os05t0513400-01)
05	25,458,856	Os05g0513450	Hypothetical protein. (Os05t0513450-00)
05	25,462,659	Os05g0513500	Similar to 50S ribosomal protein L33. (Os05t0513500-01)
05	25,464,547	Os05g0513600	Conserved hypothetical protein. (Os05t0513600-01)
05	25,467,047	Os05g0513700	Reticulon domain containing protein. (Os05t0513700-00)
05	25,474,799	Os05g0513800	Small GTP-binding protein OsRac2. (Os05t0513800-01)
05	25,474,799	Os05g0513800	Small GTP-binding protein OsRac2. (Os05t0513800-01)
05	25,477,117	Os05g0513900	Similar to Proline iminopeptidase (EC 3.4.11.5) (PIP) (Prolyl aminopeptidase) (PAP). (Os05t0513900-01)
05	25,477,117	Os05g0513900	Similar to Proline iminopeptidase (EC 3.4.11.5) (PIP) (Prolyl aminopeptidase) (PAP). (Os05t0513900-01)
05	25,489,349	Os05g0514200	OsPK4. (Os05t0514200-01)
05	25,489,525	Os05g0514250	Hypothetical protein. (Os05t0514250-00)
05	25,497,335	Os05g0514300	Similar to Tubby-like protein 3. (Os05t0514300-01) Similar to Tubby-like protein 3. (Os05t0514300-02)

05	25,497,335	Os05g0514300	Similar to Tubby-like protein 3. (Os05t0514300-01) Similar to Tubby-like protein 3. (Os05t0514300-02)
05	25,503,455	Os05g0514400	Hypothetical conserved gene. (Os05t0514400-00)

chr.	bp	gene name	gene function
08	4,105,460	Os08g0170100	NB-ARC domain containing protein. (Os08t0170100-01)
08	4,111,865	Os08g0170200	Disease resistance protein domain containing protein. (Os08t0170200-00)
08	4,116,293	Os08g0170400	Conserved hypothetical protein. (Os08t0170400-01)
08	4,128,336	Os08g0170500	Conserved hypothetical protein. (Os08t0170500-00)
08	4,135,126	Os08g0170700	Disease resistance protein domain containing protein. (Os08t0170700-00)
08	4,147,266	Os08g0170800	Conserved hypothetical protein. (Os08t0170800-01)
08	4,153,416	Os08g0170900	Winged helix repressor DNA-binding domain containing protein. (Os08t0170900-01)
08	4,153,416	Os08g0170900	Winged helix repressor DNA-binding domain containing protein. (Os08t0170900-01)
08	4,172,264	Os08g0171000	Conserved hypothetical protein. (Os08t0171000-01)
08	4,175,878	Os08g0171100	Similar to WRINKLED1 (Activator of sporamin LUC 1). (Os08t0171100-01)
08	4,177,561	Os08g0171216	Similar to Transcription factor AP2D9. (Os08t0171216-01)
08	4,179,589	Os08g0171333	Glutaredoxin domain containing protein. (Os08t0171333-00)
08	4,191,606	Os08g0171600	RNA polymerase Rpb2 domain 6 containing protein. (Os08t0171600-01)
08	4,206,406	Os08g0171700	Conserved hypothetical protein. (Os08t0171700-01)
08	4,211,443	Os08g0171800	Protein of unknown function DUF617 plant family protein. (Os08t0171800-01)
08	4,237,020	Os08g0172200	UvrB/UvrC protein domain containing protein. (Os08t0172200-01)
08	4,242,951	Os08g0172300	Similar to H1005F08.16 protein. (Os08t0172300-01) Hypothetical protein. (Os08t0172300-02)
08	4,242,951	Os08g0172300	Similar to H1005F08.16 protein. (Os08t0172300-01) Hypothetical protein. (Os08t0172300-02)
08	4,254,521	Os08g0172566	Similar to H0315A08.1 protein. (Os08t0172566-00)
08	4,268,283	Os08g0172800	Conserved hypothetical protein. (Os08t0172800-01)
08	4,290,418	Os08g0173300	Conserved hypothetical protein. (Os08t0173300-01)
08	4,300,574	Os08g0173600	Conserved hypothetical protein. (Os08t0173600-01)

08	4,303,902	Os08g0173700	Similar to Ethylene-responsive transcription factor 4 (Ethylene-responsive element binding factor 4) (Related to APETALA-2 protein 5) (EREBP-4) (AtERF4). (Os08t0173700-01)
----	-----------	--------------	---

chr	bp	gene name	gene function
9	22,564,044	Os09g0566050	Similar to Prephenate dehydratase.(Os09t0566050-00);Transcript variants=Os09t0566050-00
9	22,566,452	Os09g0566100	ENTH/VHS domaincontaining protein. (Os09t0566100-01) Hypothetical conserved gene.(Os09t0566100-0
9	22,601,494	Os09g0566550	Similar toserine/threonine protein kinase 1 CTR1. (Os09t0566550-01);Transcriptvariants=Os09t056
9	22,622,744	Os09g0566700	Protein of unknownfunction DUF1677 plant domain containing protein.(Os09t0566700-01);Transcript
9	22,631,110	Os09g0567000	Domain of unknownfunction DUF1618 domain containing protein. (Os09t0567000-01);Transcriptvariants=
9	22,644,685	Os09g0567300	Monodehydroascorbatereductase Tolerance to ROS-induced oxidative stress(Os09t0567300-01);
9	22,649,087	Os09g0567366	Similar to Aldo-ketoreductase/ oxidoreductase. (Os09t0567366-01) Similar to Aldo-ketoreductase/
9	22,654,004	Os09g0567400	Histidinephosphotransfer protein Cytokinin signaling and stress response(Os09t0567400-01)
9	22,660,725	Os09g0567500	Similar to Fatty acylcoA reductase. (Os09t0567500-01) Similar to Fatty acyl coA reductase.(Os09t
9	22,666,388	Os09g0567700	WD40 repeat-like domaincontaining protein. (Os09t0567700-01);Transcriptvariants=Os09t0567700-01;Or
9	22,677,560	Os09g0567900	Inosine/uridine-preferring nucleosidehydrolase domain containing protein. (Os09t0567900-01);Transcr
9	22,682,528	Os09g0568000	Single-stranded nucleicacid binding R3H domain containing protein. (Os09t0568000-01) SUZ domain
9	22,690,177	Os09g0568200	Similar to DNAPolymerase epsilon subunit 3. (Os09t0568200-01);Transcriptvariants=Os09t0568200-01;C
9	22,692,732	Os09g0568266	Non-protein codingtranscript. (Os09t0568266-00);Transcript variants=Os09t0568266-00
9	22,693,865	Os09g0568400	Similar to 60Sribosomal protein L40 (CEP52). (Os09t0568400-01) Similar toUbiquitin/ribosomal fus
9	22,695,120	Os09g0568500	Germin family protein.(Os09t0568500-01);Transcript variants=Os09t0568500-01;CGSNL GeneSymbol=GLP9-
9	22,697,745	Os09g0568600	Putative germin-likeprotein 9-2. (Os09t0568600-00);Transcript variants=Os09t0568600-00;CGSNL GeneS
9	22,699,925	Os09g0568700	RmlC-like jelly rollfold domain containing protein. (Os09t0568700-01);Transcriptvariants=Os09t0568
9	22,701,807	Os09g0568800	Similar to Ribosomalprotein S25 (40S ribosomal 25S subunit). (Os09t0568800-01);Transcriptvariants=
9	22,704,849	Os09g0568900	Similar to predictedprotein. (Os09t0568900-01);Transcript variants=Os09t0568900-01;Oryzabase GeneS
9	22,707,659	Os09g0569100	Haloaciddehalogenase-like hydrolase domain containing protein.(Os09t0569100-01);Transcript variant
9	22,713,282	Os09g0569200	Similar to Beta-amylase(EC 3.2.1.2) (1,4-alpha-D-glucan maltohydrolase).(Os09t0569200-01);Transc
9	22,716,748	Os09g0569300	Similar tocalmodulin-binding heat-shock protein. (Os09t0569300-01);Transcriptvariants=Os09t0569300
9	22,720,434	Os09g0569400	Beta-lactamase-likedomain containing protein. (Os09t0569400-01);Transcriptvariants=Os09t0569400-01
9	22,736,559	Os09g0569800	Serine/threonineprotein kinase-related domain containing protein.(Os09t0569800-00);Transcript vari
9	22,749,994	Os09g0570000	Serine/threonineprotein kinase-related domain containing protein.(Os09t0570000-01);Transcript vari

9	22,755,848	Os09g0570100	Protein kinase%2Ccatalytic domain domain containing protein. (Os09t0570100-01) Hypotheticalconse
9	22,758,931	Os09g0570200	Zinc finger%2CC2H2-type domain containing protein. (Os09t0570200-01) Zinc finger C2H2-type dom
9	22,762,865	Os09g0570300	Similar to Short-chaindehydrogenase Tic32. (Os09t0570300-01) Similar to Short-chaindehydrogenase
9	22,767,376	Os09g0570400	Major facilitatorsuperfamily protein. (Os09t0570400-01)Non-protein coding transcript.(Os09t0570
9	22,774,959	Os09g0570500	Zinc finger%2CRING-type domain containing protein. (Os09t0570500-01);Transcriptvariants=Os09t05705
9	22,778,763	Os09g0570600	PAP/25A core domaincontaining protein. (Os09t0570600-01)Non-protein coding transcript.(Os09t057
9	22,790,671	Os09g0570800	Isopenicillin Nsynthase family protein. (Os09t0570800-01);Transcriptvariants=Os09t0570800-01
9	22,792,228	Os09g0570850	Similar to Histone H2B.(Os09t0570850-00);Transcript variants=Os09t0570850-00
9	22,793,266	Os09g0570900	Amino acid-binding ACTdomain containing protein. (Os09t0570900-01) Similar to amino acidbinding
9	22,800,622	Os09g0571000	Protein of unknownfunction DUF966 family protein. (Os09t0571000-00); Transcriptvariants=Os09t057100
9	22,804,668	Os09g0571100	Pectin lyasefold/virulence factor domain containing protein. (Os09t0571100-01); Transcriptvariants=
9	22,806,006	Os09g0571200	C2 domain containingprotein. (Os09t0571200-01);Transcript variants=Os09t0571200-01
9	22,809,907	Os09g0571400	Cyclophilin 1.(Os09t0571400-01); Transcript variants=Os09t0571400-01;CGSNLGeneSymbol=CYP19-3;CGSNL
9	22,811,278	Os09g0571500	Lipase%2C class 3family protein. (Os09t0571500-01) Similar to lipase class 3 family protein.(Os0
9	22,821,581	Os09g0571700	Similar toEthylene-responsive transcription factor ERF096. (Os09t0571700-00);Transcriptvariants=Os
9	22,847,140	Os09g0572000	Pathogenesis-relatedtranscriptional factor and ERF domain containing protein.(Os09t0572000-01);Tra
9	22,852,331	Os09g0572100	Protein kinase%2C coredomain containing protein. (Os09t0572100-01) Protein kinase coredomain
9	22,855,317	Os09g0572200	Similar to Mutant lowphytic acid protein 1. (Os09t0572200-01) Similar to predicted protein.(Os09
9	22,883,138	Os09g0572400	Similar to Abcf2-provprotein. (Os09t0572400-01);Transcript variants=Os09t0572400-01;CGSNL GeneSymb

chr.	bp	gene name	gene function
10	22,600,684	Os10g0570001	Conserved hypothetical protein. (Os10t0570001-00)
10	22,607,642	Os10g0570200	Similar to RIR1b protein precursor. (Os10t0570200-01)
10	22,627,963	Os10g0570700	Ribosome-inactivating protein family protein. (Os10t0570700-00)
10	22,629,803	Os10g0570800	Hypothetical conserved gene. (Os10t0570800-00)
10	22,636,085	Os10g0571100	Protein of unknown function DUF498 domain containing protein. (Os10t0571100-01)
10	22,640,923	Os10g0571200	Similar to Pyruvate kinase isozyme G chloroplast (EC 2.7.1.40) (Fragment). (Os10t0571200-01)
10	22,640,923	Os10g0571200	Similar to Pyruvate kinase isozyme G chloroplast (EC 2.7.1.40) (Fragment). (Os10t0571200-01)
10	22,650,009	Os10g0571300	Similar to Protein kinase-like protein. (Os10t0571300-01)
10	22,670,628	Os10g0571600	No apical meristem (NAM) protein domain containing protein. (Os10t0571600-01)
10	22,677,765	Os10g0571900	Similar to Malic enzyme (Fragment). (Os10t0571900-00)
10	22,680,088	Os10g0572000	GTP-binding protein HSR1-related domain containing protein. (Os10t0572000-00)
10	22,682,292	Os10g0572100	Uncharacterised protein family UPF0005 domain containing protein. (Os10t0572100-01)
10	22,690,792	Os10g0572300	Leucine-rich repeat typical subtype containing protein. (Os10t0572300-01)
10	22,690,792	Os10g0572300	Leucine-rich repeat typical subtype containing protein. (Os10t0572300-01)
10	22,699,144	Os10g0572400	Conserved hypothetical protein. (Os10t0572400-01)

chr.	bp	gene name	gene function
12	5,508,787	Os12g0205500	NB-ARC domain containing protein. (Os12t0205500-01)
12	5,535,012	Os12g0205633	Conserved hypothetical protein. (Os12t0205633-00)
12	5,568,838	Os12g0206700	Hypothetical protein. (Os12t0206700-01)
12	5,569,644	Os12g0206800	Transcription factor MADS-box domain containing protein. (Os12t0206800-00)
12	5,584,743	Os12g0207000	Transcription factor. (Os12t0207000-01)
12	5,596,782	Os12g0207200	Myb/SANT-like domain domain containing protein. (Os12t0207200-01)
12	5,602,196	Os12g0207300	Similar to Clathrin coat assembly protein AP17 (Clathrin coat associated protein AP17) (Plasma membrane adaptor AP-2 17 kDa protein) (HA2 17 kDa subunit) (Clathrin assembly protein 2 small chain). (Os12t0207300-01)
12	5,612,066	Os12g0207500	Similar to ATP synthase subunit beta. (Os12t0207500-00)
12	5,614,140	Os12g0207600	Similar to ribulose-1-bisphosphate carboxylase/oxygenase large subunit. (Os12t0207600-00)
12	5,674,410	Os12g0208900	Similar to ZF-HD protein dimerisation region containing protein expressed. (Os12t0208900-00)
12	5,675,631	Os12g0208950	Hypothetical protein. (Os12t0208950-00)
12	5,685,138	Os12g0209000	Similar to sporulation protein-related. (Os12t0209000-00)
12	5,692,925	Os12g0209100	Similar to predicted protein. (Os12t0209100-01)
12	5,699,790	Os12g0209200	Similar to CONSTANS-LIKE a. (Os12t0209200-01)
12	5,703,041	Os12g0209300	Similar to ATPase AAA family protein. (Os12t0209300-00)
12	5,703,041	Os12g0209300	Similar to ATPase AAA family protein. (Os12t0209300-00)
12	5,703,041	Os12g0209300	Similar to ATPase AAA family protein. (Os12t0209300-00)
12	5,703,041	Os12g0209300	Similar to ATPase AAA family protein. (Os12t0209300-00)
12	5,703,041	Os12g0209300	Similar to ATPase AAA family protein. (Os12t0209300-00)
12	5,709,600	Os12g0209502	Hypothetical gene. (Os12t0209502-01)
12	5,736,998	Os12g0209700	Hypothetical conserved gene. (Os12t0209700-01)
12	5,740,543	Os12g0209800	Similar to NB-ARC domain containing protein expressed. (Os12t0209800-01)
12	5,749,555	Os12g0209901	Similar to NB-ARC domain containing protein expressed. (Os12t0209901-01)
12	5,752,504	Os12g0210000	Similar to NB-ARC domain containing protein expressed. (Os12t0210000-01)
12	5,756,479	Os12g0210200	Similar to Glutathione S-transferase GST 18 (EC 2.5.1.18). (Os12t0210200-01)
12	5,760,537	Os12g0210300	Similar to Glutathione S-transferase GST 18 (EC 2.5.1.18). (Os12t0210300-01)
12	5,766,036	Os12g0210400	Protein kinase%2C core domain containing protein. (Os12t0210400-01)

12	5,772,511	Os12g0210450	Hypothetical gene. (Os12t0210450-00)
12	5,775,580	Os12g0210500	Conserved hypothetical protein. (Os12t0210500-00)
12	5,786,606	Os12g0210601	Hypothetical conserved gene. (Os12t0210601-00)
12	5,801,534	Os12g0210800	Similar to 3-deoxy-D-manno-2-octulosonic acid-8-phosphate (EC 4.1.2.16) (Fragment). (Os12t0210800-01)
12	5,801,534	Os12g0210800	Similar to 3-deoxy-D-manno-2-octulosonic acid-8-phosphate (EC 4.1.2.16) (Fragment). (Os12t0210800-01)
12	5,817,169	Os12g0211000	Conserved hypothetical protein. (Os12t0211000-01)
12	5,824,939	Os12g0211100	Conserved hypothetical protein. (Os12t0211100-01)
12	5,838,245	Os12g0211301	Hypothetical conserved gene. (Os12t0211301-00)
12	5,849,738	Os12g0211400	A/G-specific adenine glycosylase MutY bacterial form domain containing protein. (Os12t0211400-00)
12	5,856,248	Os12g0211500	Leucine-rich repeat N-terminal domain containing protein. (Os12t0211500-01)
12	5,857,510	Os12g0211600	Hypothetical protein. (Os12t0211600-01)
12	5,866,457	Os12g0211900	Hypothetical conserved gene. (Os12t0211900-01)
12	5,892,605	Os12g0212100	Similar to A/G-specific adenine DNA glycosylase. (Os12t0212100-00)
12	5,905,934	Os12g0212400	Hypothetical protein. (Os12t0212400-01)
12	5,959,962	Os12g0214700	Conserved hypothetical protein. (Os12t0214700-01)
12	6,059,272	Os12g0215100	Hypothetical conserved gene. (Os12t0215100-00)
12	6,130,297	Os12g0215800	Hypothetical protein. (Os12t0215800-00)
12	6,138,572	Os12g0215950	Similar to Leucine Rich Repeat family protein%2C expressed. (Os12t0215950-01)
12	6,140,537	Os12g0216000	Hypothetical protein. (Os12t0216000-01)
12	6,168,271	Os12g0216500	Hypothetical gene. (Os12t0216500-01)
12	6,224,005	Os12g0217300	Hypothetical protein. (Os12t0217300-01)Hypothetical gene. (Os12t0217300-02)
12	6,224,005	Os12g0217300	Hypothetical protein. (Os12t0217300-01)Hypothetical gene. (Os12t0217300-02)
12	6,224,187	Os12g0217400	Similar to Leucine Rich Repeat family protein expressed. (Os12t0217400-00)
12	6,259,242	Os12g0217800	Hypothetical gene. (Os12t0217800-01)
12	6,301,536	Os12g0218100	Similar to cDNA clone: J065028F04 full insert sequence. (Os12t0218100-01)
12	6,324,997	Os12g0218300	Conserved hypothetical protein. (Os12t0218300-01)
12	6,343,046	Os12g0218500	Hypothetical conserved gene. (Os12t0218500-01)

12	6,344,596	Os12g0218600	Hypothetical conserved gene. (Os12t0218600-01) Hypothetical gene. (Os12t0218600-02)
12	6,344,596	Os12g0218600	Hypothetical conserved gene. (Os12t0218600-01) Hypothetical gene. (Os12t0218600-02)
12	6,354,156	Os12g0218800	Hypothetical protein. (Os12t0218800-01)
12	6,425,710	Os12g0219700	Hypothetical protein. (Os12t0219700-01)
12	6,437,460	Os12g0220001	Conserved hypothetical protein. (Os12t0220001-01)
12	6,489,531	Os12g0220800	Hypothetical protein. (Os12t0220800-01)
12	6,489,582	Os12g0220833	Hypothetical genes. (Os12t0220833-00)
12	6,506,700	Os12g0220866	Hypothetical gene. (Os12t0220866-00)
12	6,563,411	Os12g0221400	Similar to cDNA clone:001-112-B09 full insert sequence. (Os12t0221400-01)
12	6,576,372	Os12g0221600	Conserved hypothetical protein. (Os12t0221600-01)
12	6,585,435	Os12g0221700	Hypothetical conserved gene. (Os12t0221700-00)
12	6,625,464	Os12g0222300	Conserved hypothetical protein. (Os12t0222300-01)
12	6,643,580	Os12g0222650	Hypothetical gene. (Os12t0222650-01)
12	6,647,365	Os12g0222800	Leucine-rich repeat N-terminal domain containing protein. (Os12t0222800-00)
12	6,650,911	Os12g0222900	Leucine-rich repeat N-terminal domain containing protein. (Os12t0222900-01)
12	6,652,881	Os12g0223000	Hypothetical protein. (Os12t0223000-01)
12	6,677,664	Os12g0223300	Similar to Outer membrane cytochrome b(5) (Fragment). (Os12t0223300-01)
12	6,677,664	Os12g0223300	Similar to Outer membrane cytochrome b(5) (Fragment). (Os12t0223300-01)
12	6,715,421	Os12g0223700	Conserved hypothetical protein. (Os12t0223700-01)
12	6,729,703	Os12g0224000	Similar to Diacylglycerol kinase 1 (EC 2.7.1.107) (Diglyceride kinase 1) (DGK 1) (DAG kinase 1). (Os12t0224000-01)
12	6,754,961	Os12g0224400	Similar to Exostosin family protein expressed. (Os12t0224400-01)
12	6,760,060	Os12g0224500	Zinc finger PHD-type domain containing protein. (Os12t0224500-01)
12	6,799,607	Os12g0225100	Nucleotide-binding alpha-beta plait domain containing protein. (Os12t0225100-01)
12	6,799,607	Os12g0225100	Nucleotide-binding alpha-beta plait domain containing protein. (Os12t0225100-01)
12	6,799,840	Os12g0225150	Hypothetical protein. (Os12t0225150-00)
12	6,811,122	Os12g0225200	Conserved hypothetical protein. (Os12t0225200-01)

12	6,815,299	Os12g0225300	Hypothetical conserved gene. (Os12t0225300-01)
12	6,815,299	Os12g0225300	Hypothetical conserved gene. (Os12t0225300-01)
12	6,815,299	Os12g0225300	Hypothetical conserved gene. (Os12t0225300-01)
12	6,865,308	Os12g0225700	Conserved hypothetical protein. (Os12t0225700-01)
12	6,869,987	Os12g0225900	Allyl alcohol dehydrogenase. (Os12t0225900-01)
12	6,889,996	Os12g0226400	Similar to 2-alkenal reductase. (Os12t0226400-00)
12	6,895,550	Os12g0226500	Hypothetical protein. (Os12t0226500-00)
12	6,911,276	Os12g0226700	Similar to Allyl alcohol dehydrogenase. (Os12t0226700-01)
12	6,916,185	Os12g0226900	Similar to Allyl alcohol dehydrogenase. (Os12t0226900-01)
12	6,916,214	Os12g0227050	Hypothetical protein. (Os12t0227050-00)
12	6,932,587	Os12g0227400	Allyl alcohol dehydrogenase. (Os12t0227400-01)
12	6,945,206	Os12g0227500	Similar to Beta-glucosidase aggregating factor. (Os12t0227500-00)
12	6,945,292	Os12g0227766	Hypothetical protein. (Os12t0227766-00)
12	6,962,407	Os12g0228032	Hypothetical protein. (Os12t0228032-00)
12	6,966,725	Os12g0228300	Conserved hypothetical protein. (Os12t0228300-01)
12	6,973,106	Os12g0228325	Hypothetical protein. (Os12t0228325-00)
12	6,973,270	Os12g0228350	Conserved hypothetical protein. (Os12t0228350-00)
12	6,976,909	Os12g0228400	Similar to cDNA clone:J033003J10 full insert sequence. (Os12t0228400-01), Similar to cDNA clone:J033003J10 full insert sequence. (Os12t0228400-02)
12	6,976,909	Os12g0228400	Similar to cDNA clone:J033003J10 2C full insert sequence. (Os12t0228400-01)
12	6,982,814	Os12g0228450	Conserved hypothetical protein. (Os12t0228450-00)
12	6,985,010	Os12g0228500	Hypothetical conserved gene. (Os12t0228500-01)
12	7,011,262	Os12g0228650	Hypothetical protein. (Os12t0228650-00)
12	7,011,412	Os12g0228700	Similar to 32 kDa protein. (Os12t0228700-00)
12	7,014,180	Os12g0228800	Similar to Calmodulin-like protein. (Os12t0228800-01)
12	7,017,361	Os12g0228900	Similar to Leucine Rich Repeat family protein%2C expressed. (Os12t0228900-01)
12	7,028,694	Os12g0229100	Hypothetical conserved gene. (Os12t0229100-01)
12	7,066,143	Os12g0229600	Ankyrin repeat domain containing protein. (Os12t0229600-00)
12	7,095,860	Os12g0230100	Similar to ATP-dependent Clp protease ATP-binding subunit clpA homolog
12	7,103,232	Os12g0230133	Conserved hypothetical protein. (Os12t0230133-01)
12	7,111,787	Os12g0230200	Similar to Protein kinase-like protein. (Os12t0230200-01)

12	7,111,787	Os12g0230200	Similar to Protein kinase-like protein. (Os12t0230200-01)
12	7,112,683	Os12g0230233	Hypothetical protein. (Os12t0230233-00)
12	7,115,776	Os12g0230266	Hypothetical genes. (Os12t0230266-00)
12	7,130,738	Os12g0230600	Protein of unknown function DUF1685 family protein. (Os12t0230600-01)
12	7,165,808	Os12g0231000	TGF-beta receptor type I/II extracellular region family protein. (Os12t0231000-01)
12	7,165,808	Os12g0231000	TGF-beta receptor type I/II extracellular region family protein. (Os12t0231000-01)
12	7,171,810	Os12g0231100	Conserved hypothetical protein. (Os12t0231100-01)
